INSTITUTO IMDEA NANOCIENCIA

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

TESIS DOCTORAL – PhD THESIS:

The folding complexity of TERRA G-quadruplex unveiled at the single-molecule level

Irene Gutiérrez Pérez

Memoria presentada para optar al título de Doctor
en Física de la Luz y de la Materia

Septiembre, 2017

Director: Dr. José Ricardo Arias González de la Aleja
Optical Nanomanipulation Lab
Instituto IMDEA Nanociencia

Contents

Resumen

Los telómeros son estructuras nucleoproteicas que se encuentran en los extremos de los cromosomas lineales para protegerlos frente a la degradación. Dentro del genoma humano, son considerados como los poseedores de la mayor cantidad de secuencias nucleicas propensas a formar estructuras tipo *G-quadruplex*, debido a sus extremos 3' de cadena sencilla ricos en guaninas (Gs).

Las estructuras tipo *G-quadruplex* se forman por el apareamiento, mediante enlaces tipo *Hoogsteen*, de cuatro hebras de ADN o ARN con una alta concentración de Gs. Concretamente, surgen del apilamiento de dos o más tétradas de Gs estabilizadas mediante cationes monovalentes, típicamente sodio o potasio. La demostración de su existencia *in vivo* en los telómeros sugiere que dichas estructuras no canónicas podrían desempeñar una función protectora en los mismos y que, por consiguiente, estarían involucradas en la estabilidad cromosómica.

Recientemente, se ha demostrado que los telómeros pueden ser transcritos, desde las regiones subteloméricas hacia los extremos de los cromosomas, en moléculas *TERRA* (del inglés, *TElomeric Repeat-containing RNA*). Al igual que sus equivalentes en ADN, estos tránscritos no codificantes podrían plegarse formando tándems de *G-quadruplexes* y desempeñar funciones relevantes a nivel celular, tales como la regulación de la longitud de los telómeros, la inhibición de la actividad enzimática de la telomerasa, la formación de la heterocromatina telomérica y/o la protección de los telómeros.

En esta tesis se ha investigado la estabilidad mecano-química de moléculas *TERRA* mediante las técnicas de microscopía de fuerzas atómicas (AFM; del inglés, *Atomic Force Microscopy*) —en modo imagen— y pinzas ópticas (OT; del inglés, *Optical Tweezers*) —como técnica de espectroscopía de fuerzas a nivel de molécula individual—.

Mediante AFM, se ha logrado, por primera vez, la visualización directa de moléculas largas de TERRA en aire (*Capítulo 2*). Para ello, se ha desarrollado un protocolo bioquímico, riguroso

y eficaz, que permite la comparación *in situ* de estructuras de ácidos nucleicos tipo cuádruplex y dúplex por medio de este tipo de microscopía. Además, a través de las imágenes obtenidas, se ha conseguido dilucidar la configuración adoptada por dichas moléculas sobre una superficie, lo cual ha resultado consistente con lo observado previamente para *G-quadruplexes* de DNA.

Mediante el uso de pinzas ópticas, se ha estudiado su dinámica de plegamiento, revelándose una inesperada complejidad estructural inherente a las especies de ARN (*Capítulo 3*). Concretamente, se han analizado moléculas de ARN que contienen entre cinco y siete repeticiones de la secuencia telomérica GGGUUA; es decir, secuencias que sólo pueden dar lugar a un único *G-quadruplex*. De acuerdo a los resultados obtenidos, el plegamiento de las moléculas *TERRA* en forma de *G-quadruplexes* se encontraría limitado por la capacidad inherente al ARN de auto-plegarse sobre sí mismo de manera aleatoria, formando así estructuras más favorables a nivel entrópico. Este bloqueo estocástico confiere a estas moléculas un patrón de plegamiento bimodal, el cual no se ha encontrado en sus análogos en ADN.

El bloqueo estocástico que sufren los *G-quadruplexes* de TERRA debido a la condensación del ARN puede ser relevante tanto para el entendimiento de sus posibles funciones reguladoras dentro de la célula, como en el diseño de fármacos con unión selectiva a dichas estructuras cuaternarias para el tratamiento de diversas enfermedades, incluyéndose entre ellas el cáncer.

Abstract

Telomeres are nucleoprotein structures located at the end of linear chromosomes that are essential to protect them from degradation. In the human genome, they concentrate the highest amount of potential G-quadruplex-forming sequences due to their 3'-end G-rich single-stranded DNA overhangs.

A G-quadruplex (GQ) is a four-stranded structure that arises from the stacking of two or more planar arrangements of four guanines (*G-tetrads*) stabilized by Hoogsteen hydrogen bonds and monovalent cations, typically sodium or potassium. Telomeric GQ visualization *in vivo* suggests that these non-canonical structures would have a role in telomere protection and, therefore, in chromosome stability.

Recently, it has been found that telomeres can be transcribed into *TElomeric Repeat-containing RNA* (*TERRA*) molecules. As their DNA counterparts, these non-coding transcripts are susceptible to fold into multiple GQ blocks and could perform various cellular functions, such as regulation of telomere length, inhibition of telomerase, telomeric heterochromatin formation and telomere protection.

In this thesis, TERRA molecules have been investigated at the single-molecule level by atomic force microscopy (AFM), in the imaging mode, and optical tweezers (OT), as a force-spectroscopic tool, to understand their mechano-chemical stability.

By AFM, a direct visualization of long TERRA molecules in ambient air has been obtained for the first time (*Chapter 2*). In this regard, a rigorous and efficient biochemical protocol that allows an *in situ* comparison of quadruplex and duplex nucleic acids structures has been developed. The resulting images have afforded the elucidation of the assembly mode of TERRA GQ tandems over a surface, finding consistency with earlier images of DNA GQs.

By OT, the folding dynamics of TERRA molecules has been characterized revealing an unexpected structural complexity inherent to the RNA species (*Chapter 3*). More in depth, RNA

molecules with five to seven repeats of the G-rich TERRA sequence GGGUUA, thus capable of forming only one GQ, have been analysed finding that the potential of RNA telomeric sequences to fold into a GQ conformation is limited by the intrinsic capacity of RNA to self-associate randomly and further condense into entropically more favourable structures. This behaviour confers on these short molecules a unique bistable-folding capacity that is not found in their DNA analogues.

On balance, the stochastic blockage of TERRA GQ assembly generated by RNA condensation may show relevant to understand the potential regulatory functions of TERRA molecules inside the cell, and in the design of GQ-binding drugs for the treatment of several human diseases, including cancer.

List of Abbreviations

A	Adenine
AFM	Atomic Force Microscopy
bp	Base-pair
C	Cytosine
DNA	Deoxyribonucleic acid
ds	Double-stranded
dsDNA	Double-stranded DNA
dsRNA	Double-stranded RNA
G	Guanine
GQ	G-quadruplex
mT	Minitweezers
MFP	Molecular Force Probe
MT	Magnetic Tweezers
nt	Nucleotide
OT	Optical Tweezers
PDB	Protein Data Bank
PG4s	Potential G-quadruplex-forming sequences
PSD	Position-Sensitive Detector
RNA	Ribonucleic acid
SMT	Single Molecule Technique
ss	Single-stranded
ssDNA	Single-stranded DNA
ssRNA	Single-stranded RNA
T	Thymine
TERRA	Telomeric Repeat-containing RNA
TERRA-16	TERRA molecules with 16 GGGUUA repeats
TERRA-25	TERRA molecules with 25 GGGUUA repeats
U	Uracil
UTR	Untranslated Region

Chapter 1

Introduction

1.1. Fundamentals of nucleic acids

Nucleic acids are key biological macromolecules in the origin and maintenance of life. In particular, they are responsible for storing, transmitting and expressing genetic information from one generation to the next. The structure of every protein, and ultimately of every biomolecule and cellular component, is a product of the information programmed into a specific nucleic acid sequence.

Two types of this material basis of genetic inheritance are distinguished in nature: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Although both of them are capable of carrying and preserving information using a four-letter alphabet, DNA was the selected molecule by most organisms, leaving RNA confined to some viral genomes. Over the past few years, RNA has emerged as a far more relevant molecule than previously expected thanks to the discovery of RNA interference [1-3].

Beyond their biological relevance, DNA and RNA have also become powerful platforms for nanotechnological applications because of their exceptional structural and functional diversity: from compatible biomaterials or new treatments for several genetic diseases to potential alternatives to current data storage systems [4-13].

This section provides an overview of the chemical nature of nucleic acids as well as a detailed examination of their structures and functions.

1.1.1. The building blocks of nucleic acids

Nucleic acids, DNA and RNA, are formed by the polymeric arrangement of four different subunits, called nucleotides, covalently connected through phosphodiester bonds. These

monomeric subunits could be further divided into phosphate groups and four distinct nucleosides. The latter were subsequently identified as consisting of a pentose sugar and one of four nitrogen-containing heterocyclic bases. Thus, each repeating unit in a nucleic acid polymer comprises these three characteristic components linked together: a phosphate group, a pentose and a nitrogenous base (*Fig.1.1*).

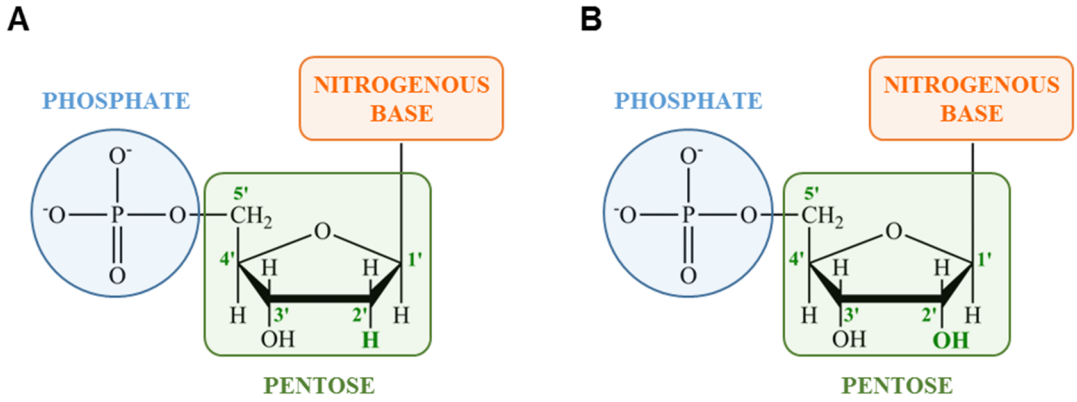

FIGURE 1.1. Structure of nucleotides. General structure of a **(A)** deoxyribonucleotide and **(B)** ribonucleotide, showing their three units: phosphate group, nitrogenous base and pentose ring with its numbering convention.

Pentoses are cyclic monosaccharides of five carbon atoms. Nucleic acids contain two kinds of them: *2'-deoxy-D-ribose* in the case of DNA and *D-ribose* for ribonucleotide units of RNA. Both types of pentoses are in their β-furanose form (closed five-membered ring) and their carbon atoms are numbered as 1', 2', 3'… to distinguish them from those of the nitrogenous bases (*Fig. 1.2*).

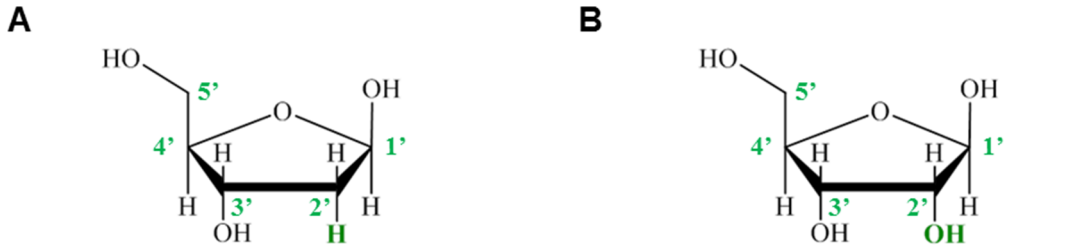

FIGURE 1.2. Types of pentoses in nucleic acids. **(A)** *2'-deoxy-D-ribose,* in the case of DNA and **(B)** *D-ribose,* for ribonucleotide units. Numbering convention is also specified.

The nitrogenous bases are planar aromatic heterocyclic compounds derived from pyrimidine (based on a single aromatic ring) or purine (based on two connected aromatic rings). Both DNA and RNA contain two major purine bases, adenine and guanine, and two major

pyrimidines. In both nucleic acids, one of the pyrimidines is cytosine, but the second one is not the same: it is thymine in DNA and uracil in RNA. Their structures are shown in *Figure 1.3*.

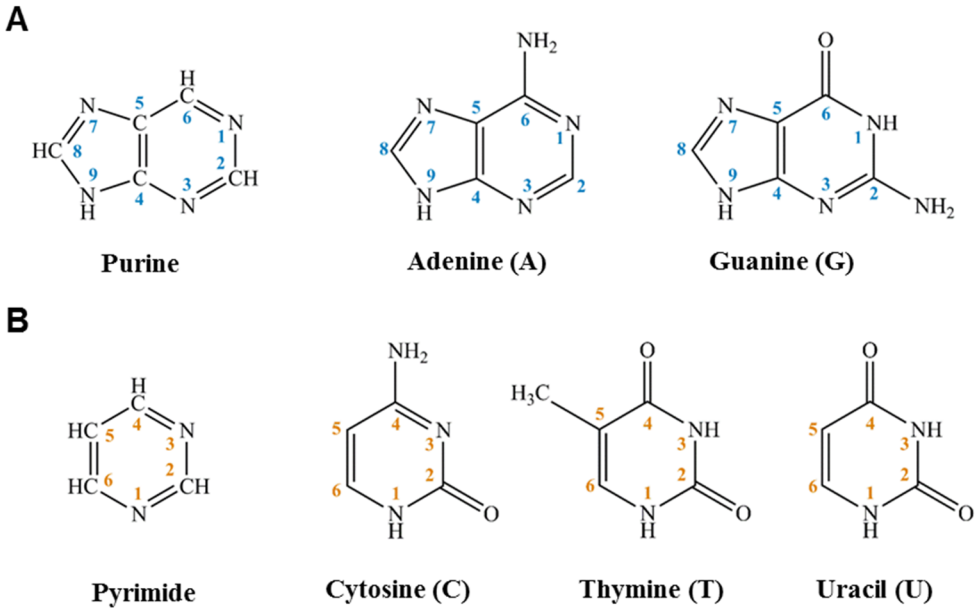

FIGURE 1.3. Major forms of nitrogenous bases in nucleic acids. **(A)** Purine and **(B)** pyrimidine-derived bases showing the numbering convention.

The base of each nucleotide is covalently linked to the sugar through a β-N-glycosidic bond. β means that the base is above the plane of the sugar when viewed onto the plane and, therefore, on the same face of the plane as the 5' hydroxyl substituent. In pyrimidine bases, this bond is created between *N-1* and the sugar *C-1'* atom whereas in the purines, *N-9* is the nitrogen atom involved. The formation of this linkage provokes the release of a water molecule, which is generated from a hydroxyl group of the pentose and a hydrogen atom of the base (*Fig. 1.4*).

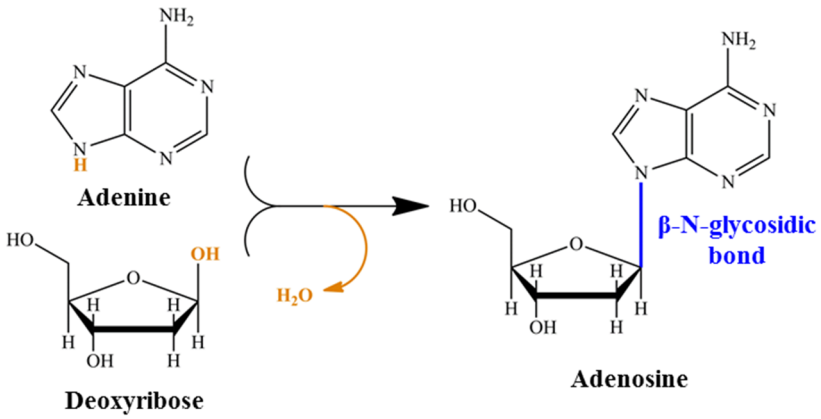

FIGURE 1.4. Formation of β-N- *glycosidic* linkage between a nitrogenous base and a pentose. Here, it is shown how an adenine bonds to a deoxyribose by a β-N- *glycosidic* linkage, releasing a water molecule.

The successive nucleotides of DNA and RNA are covalently linked by phosphodiester bonds between the 5'-hydroxyl group of a pentose and the 3'-hydroxyl group of the following, conferring to the nucleic acid chains a specific polarity and differentiated 5' and 3' ends. By definition, the 5’ end lacks a nucleotide at the 5’ end position and the 3’ end, at the 3’ position. Other groups (most often one or more phosphates) may be present on one or both ends. Therefore, the covalent chains of nucleic acids could be described as a backbone of alternating phosphate and pentose residues with different nitrogenous bases attached at regular intervals, giving rise to a particular sequence in which the genetic information is encoded (*Fig. 1.5*).

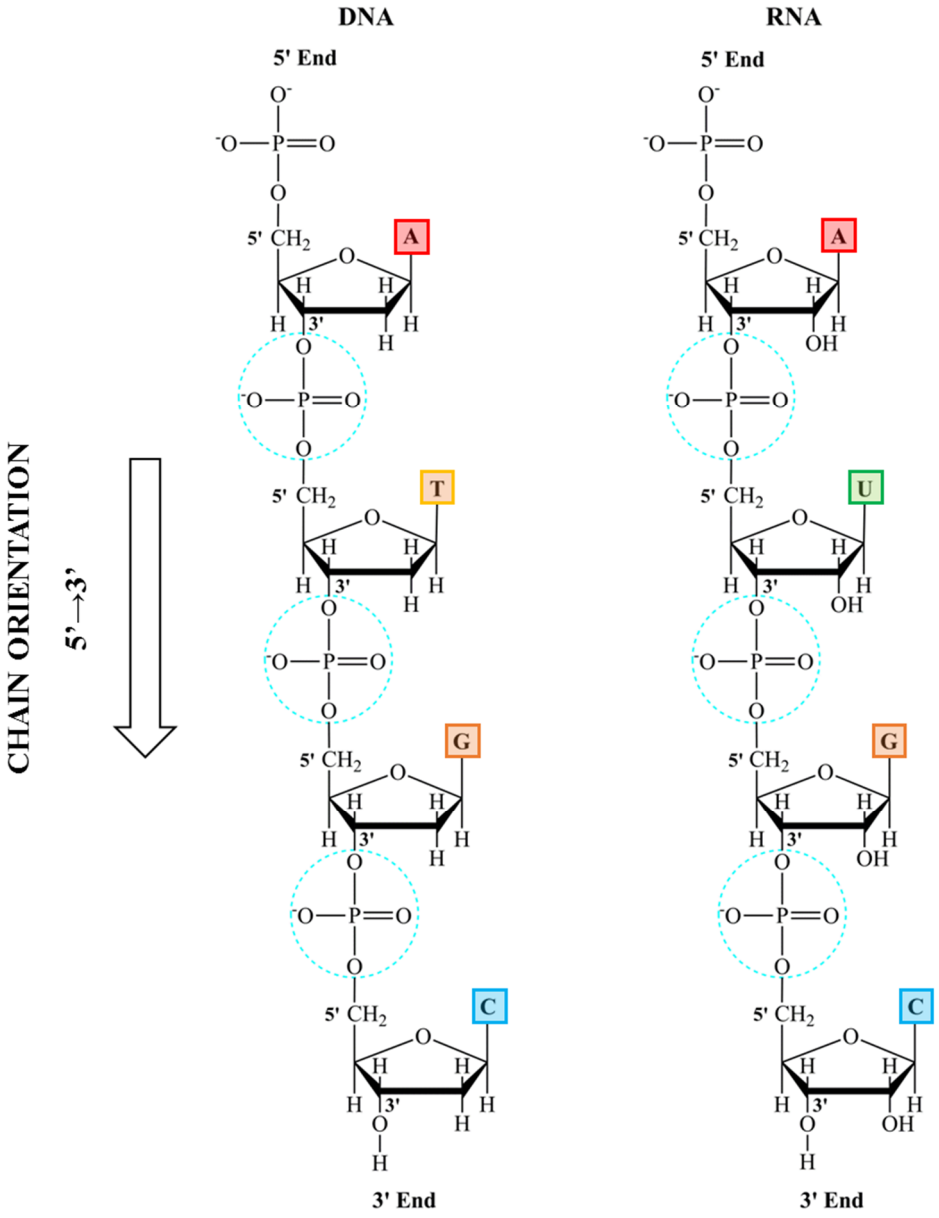

FIGURE 1.5. Phosphodiester linkages in the covalent backbone of DNA and RNA. The phosphodiester bonds linking successive nucleotide units are highlighted with light blue, dashed lines.

1.1.2. The Watson-Crick model of DNA

Under the name of "*nuclein*", DNA was firstly isolated and characterized from leukocyte cell nuclei by Friedrich Miescher in 1868. In the 1930s, Albrecht Kossel and Phoebus Levene established that nuclein was the deoxyribonucleic acid, defined as a molecule with different repetitive subunits. However, its function as genetic information carrier was not discovered until the 1940s: firstly, with the work of Oswald T. Avery, Colin MacLeod and Maclyn McCarty in 1944 and, then, in 1952, with the experiments by Alfred D. Hershey and Martha Chase [14, 15].

At the same time, hard efforts were being done to elucidate the molecular structure of DNA. Erwin Chargaff and his colleagues found that in DNA the molar ratios of adenine:thymine and cytosine:guanine were both unity [16]. Meanwhile, Rosalind Franklin and Maurice Wilkins unravelled the helical structure of DNA using X-ray diffraction [17]. In 1953, taking into account all of these findings, James Watson and Francis Crick postulated a three-dimensional model of the DNA structure [18].

The **Watson-Crick model of DNA**, also referred to as B-form DNA or B-DNA, consist of two antiparallel helical DNA chains wound around the same axis to form a right-handed double helix (*Fig.1.6*). The hydrophilic backbones of alternating deoxyribose and phosphate groups are on the outside of the double helix, facing the surrounding water, whereas the hydrophobic purine and pyrimidine bases of both strands remain stacked inside the double helix due to a combination of van der Waals and dipole-dipole interactions between them. In addition to the base-stacking interactions, the DNA double helix maintains a regular structure thanks to hydrogen bonding between complementary base pairs: A=T, connected through two hydrogen bonds, and G≡C, through three. Another characteristic features are the opposite and unequally sized grooves that run continuously along the entire length of the molecule chain. So-called *minor* and *major grooves*, they appear because of the offset pairing of the two DNA strands.

Regarding the periodicities observed in the X-ray diffraction patterns of DNA fibres by Franklin and Wilkins, Watson and Crick proposed a structure in which the vertically stacked bases inside the double helix would be 3.4 Å apart and justified the secondary repeat distance of about 34 Å in the presence of 10 base pairs in each complete turn of the double helix. However, subsequent measurements in aqueous solution, revealed 10.5 base pairs per helical turn [19].

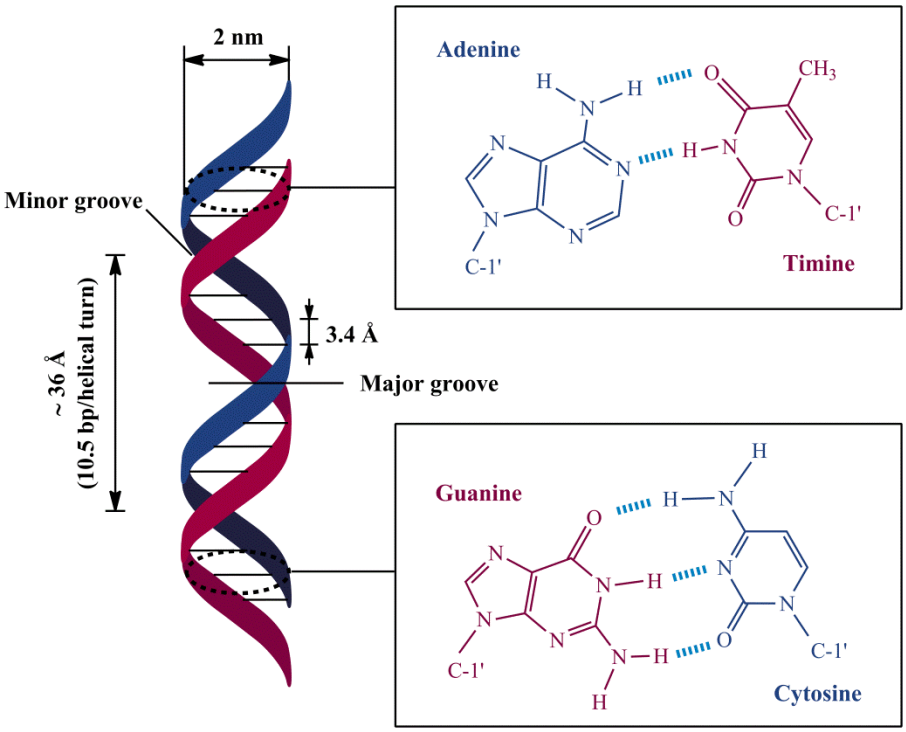

FIGURE 1.6. Watson-Crick model for the structure of DNA. Schematic representation of a double DNA helix, showing its dimensions. The offset pairing of the two strands creates two types of grooves on the surface of the duplex, whose names reflect their differences in size. Right insets represent the hydrogen-bonding patterns (light blue dashed lines) between the complementary base pairs A=T and G≡C.

1.1.3. Non-B DNA structures

B-DNA is the most stable structure for a random sequence DNA molecule under physiological conditions. However, structural transitions into other DNA forms can occur induced by changes in environmental conditions, protein binding or superhelical tension.

A-form appears when the relative humidity is reduced to less than about 75%. As B-DNA, it is a right-handed double helix but wider and shorter. It measures 2.3 nm wide and has 11 base pairs per turn. Moreover, its base pairs are tilted about 20º to the helix axis rather than being perpendicular as in B-form. These structural changes deepen the major groove while making the minor groove shallower [20-23]. Biochemical, crystallographic and computer simulation analyses of the A-DNA structure as well as protein-DNA complexes indicate that this conformation may either form upon binding of certain proteins to DNA, or be an important intermediate step in forming the strongly distorted DNA conformation observed within at least some complexes with proteins [24-26]. Recently, single-molecule experiments have proved that the A-form is not an intrinsic conformation of dsDNA, unlike in dsRNA (see next subsection).

In particular, these experiments have found no cooperative B-A transition, which should be promoted by a bistability of the molecule itself [27-29]. A contour-length reduction compatible with the so-called A-DNA may be then stabilized during protein binding or by the DNA condensation involved at low humidity in bulk experiments [30].

Z-form, whose name is due to the zigzag appearance of the backbone, is adopted by DNA segments with alternating sequence of purine and pyrimidine bases, especially C and G or 5-methyl-C and G residues. This double helix is longer and narrower than B-form; it measures 3.8 nm wide and has 12 base pairs per turn and, unlike A and B-forms, is left-handed. High salt concentrations are required to minimize electrostatic repulsion between the backbone phosphates, which are closer to each other than in A- and B-DNA. These Z-DNA tracts may play a role in regulating the expression of some genes or in genetic recombination, but it is still under investigation [31].

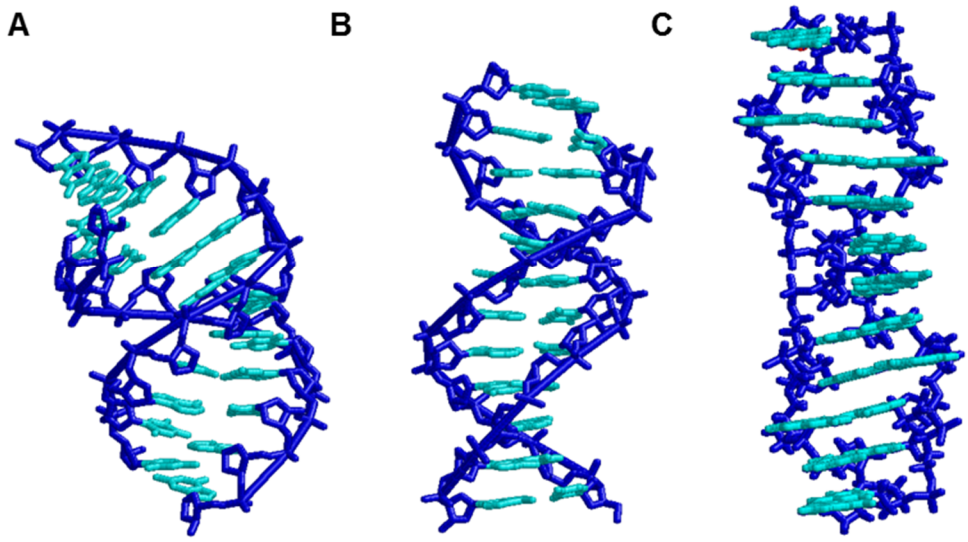

FIGURE 1.7. Comparison of A, B and Z-forms of DNA. (A) Crystal structure of an A-DNA dodecamer featuring an alternating pyrimidine-purine sequence (PDB ID: 5MVT) [32]. **(B)** Structure of a B-DNA dodecamer, d(CpGpCpGpApApTpTpCpGpCpG) (PDB ID: 1BDNA) [33]. **(C)** Structure of a Z-DNA dodecamer, generated using a computer program from standard coordinates [34]. Nitrogenous bases are represented in light blue and the backbone of alternating phosphate and pentose residues in dark blue.

In addition to Z-form, another non-B DNA structures related to rare sequences have been identified. Hairpins, cruciforms, triplex or quadruplex structures are some examples of that.

Hairpins and cruciforms, also known as stem-loop structures, arise from palindromic DNA regions; that is, DNA fragments with complementary inverted repeats (*Fig. 1.8*). In the

formation of a hairpin, only a single DNA strand is needed so it can be produced during replication, bacterial conjugation, natural transformation or viral infections. To create a cruciform, two palindromic strands are required and its stability is enhanced by DNA supercoiling. This latter stem-loop structure is fundamentally important for a wide range of gene biological processes, including replication, regulation of gene expression, nucleosome structure and recombination. Moreover, it is implicated in the evolution and development of some diseases, such as cancer and Werner's syndrome [35, 36].

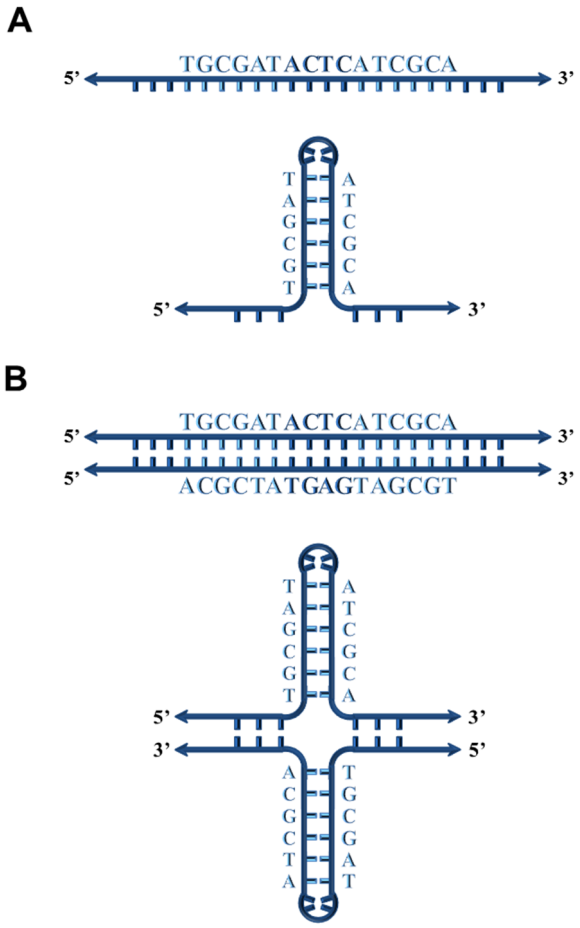

FIGURE 1.8. Schematic representation of hairpin and cruciform structures. Palindromic DNA (or RNA) sequences can form unusual structures with intrastrand base pairing: **(A)** hairpins, when only a single DNA strand is involved or **(B)** cruciform, if both strands participate.

Other non-canonical DNA structures involve three or even four DNA strands. Nucleotides participating in a Watson-Crick base pair (A=T, G≡C) can form additional hydrogen bonds, particularly with functional groups arrayed in the major groove, giving rise to *Hoogsteen-type base pairing*. This special type of hydrogen bonding allows pairing between, for example, a thymine with an A=T pair, T=A*T, or a cytidine residue (if protonated) with a G≡C nucleotide

pair, C≡G*C$^+$ (*Fig. 1.9 (A)*). **Triple-stranded (*triplex*) DNA structures**; that is, B-form double helices with a third strand coupled in their major grooves, are generated using this non-Watson-Crick base-pairing [37-40]. Four DNA strands can also pair in that way forming a **quadruplex**, but this occurs readily only for sequences with a very high proportion of guanosine residues at physiological conditions (see *Section 2* for a deeper insight). Cytosine-rich DNA sequences are able to self-fold into tetraplex structures, named ***i-motif***, at slightly acidic conditions [41, 42].

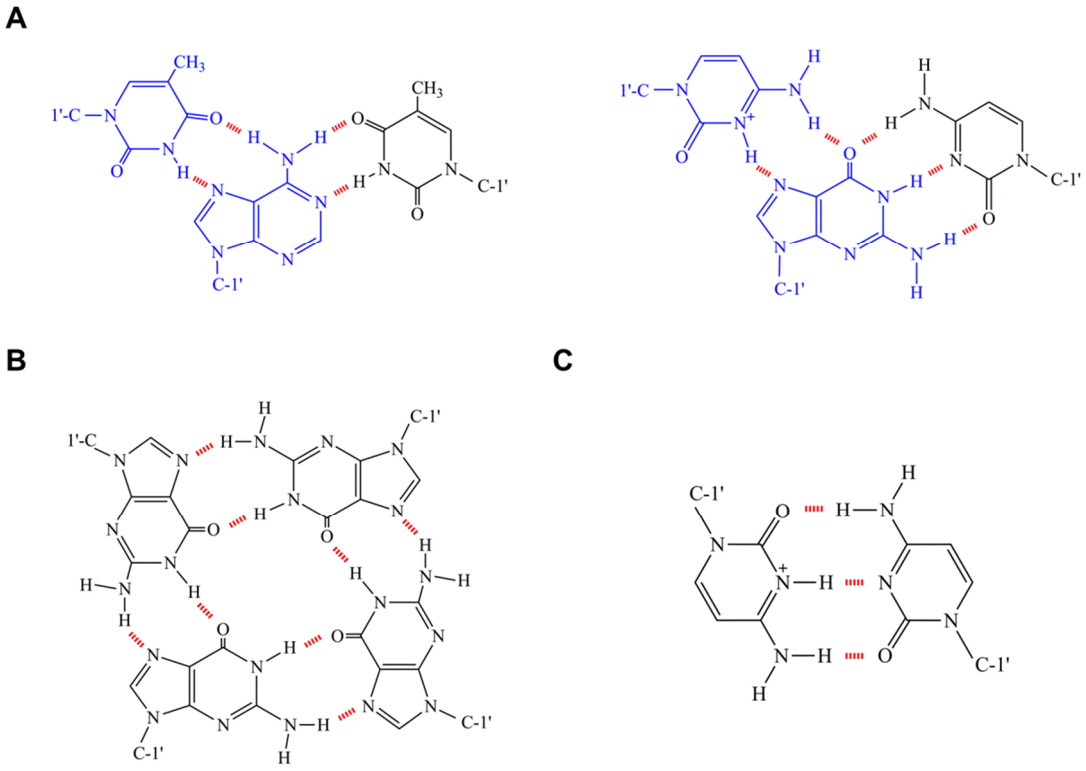

FIGURE 1.9. Hoogsteen-type base pairing. Common base-pairing patterns in: **(A)** a DNA triplex, T=A*T (left) and C≡G*C$^+$ (right); **(B)** a guanosine tetraplex, and (C) in an i-motif arrangement structure, C*C$^+$. Hydrogen-bonds are represented in red dashed lines.

1.1.4. Structural and functional diversity of RNA molecules

The relationship between structure and function in RNA remains unclear. More and more scientists are aware of the importance of elucidating this issue, considering it as the next step to be solved on the road to understand the molecular basis of life. Here, a brief review about the structural complexity of RNA and their large number of functional roles is presented.

1.1.4.1. Three-dimensional RNA structures

All RNA molecules, except the RNA genomes of certain viruses, are produced during transcription from information permanently stored in DNA. Basically, an enzyme system converts the genetic information in a segment of double-stranded DNA into a single-stranded RNA molecule with a base sequence complementary to one of the DNA chains. Although only differing in that RNA has a hydroxyl group at the 2' position of the pentose, and uracil instead of thymine (see *subsection 1.1.1*), the structural diversity of RNA is much greater than in DNA. It is due to the single-stranded nature they adopt when performing most of its functions.

Single-stranded RNA molecules tend to assume a right-handed helical conformation dominated by base-stacking interactions, which are stronger between two purines than between a purine and pyrimidine or between two pyrimidines. Nevertheless, more complex RNA structures are found in nature, from hairpin loops created by self-complementary regions of a single strand (see *Fig.1.8*) to structures that involve two, three or four strands.

Double-stranded RNA structures, either hybrids or not [43, 44], follow the same base pairing pattern as DNA (G≡C and A=U, or with occasional T residues in some RNAs), but also including the interaction between G and U residues when complementary sequences in both strands pair with each other. If both strands pair perfectly, they adopt an A-form right handed conformation (B-form has not been observed and Z-form has only been obtained synthetically under very high salt or high temperature conditions). However, perfectly paired RNA strands are uncommon over long sequences. Breaks in a regular A-form helix, caused by mismatched or unmatched bases in one or both strands, are common and result in bulges or internal loops. RNA duplexes can also be formed by ribose zipper motifs; that is, consecutive hydrogen-bonding interactions between ribose 2′-hydroxyls groups of each strand.

Recent experiments by optical and magnetic tweezers, besides AFM, have confirmed on a one-by-one basis that dsRNA is an intrinsic A-form conformation [45] in contrast to DNA, whose A-like conformation has only been observed in bulk experiments where aggregation cannot be prevented [27-30].

Regarding other RNA complex structures, **triplexes** are found in almost all large structured RNAs [46]. Mainly, they consist of double RNA helices with a third strand coupled in their minor grooves. In contrast, **RNA quadruplexes** follows the same folding pattern as their DNA

counterparts, being found in guanine rich regions of telomeres and mRNAs [47-52]. A deeper insight into this special structure can be found in *Section 1.2* of this chapter.

1.1.4.2. Functions of RNA structures

A variety of RNA forms have been distinguished based on its diverse and often complex biological functions, including *messenger RNA*, *transfer RNA*, and *ribosomal RNA*.

Messenger RNA (mRNA) is essentially a copy of a section of DNA, commonly called transcript, which carries to ribosomes the necessary genetic information to synthesise the amino acid sequence of a functional protein. *Transfer RNA (tRNA)*, folded into a cloverleaf-shaped structure, serves as molecular adaptor in protein synthesis, carrying amino acids to the ribosomes and arranging them according to the nucleotide sequence of the mRNA. *Ribosomal RNA (rRNA)*, together with a set of basic proteins, forms the ribosomes, the cellular organelles where protein synthesis, i.e. translation, takes place.

There are still more RNA forms. For instance, short RNAs are not only part of organelles like ribosomes and spliceosomes, but also of some enzymes like the RNA polymerases. Other classes of RNA species include *microRNAs (miRNAs)*, *small interfering RNAs (siRNAs)*, and *piwi-interacting RNA (piRNAs)*, all of which perform important functions in the cell but are not translated into proteins. The discovery of these RNAs has been one of the most exciting advances in recent years, and there is currently a lot of interest in the use of these molecules for therapies [53-59].

1.2. G-quadruplexes

In spite of the early breakthrough of G-quadruplexes in 1962 [60], it has not been until the last two decades when the interest in this four-stranded helical structures has increased. They had been considered as structural curiosities observed *in vitro* [61, 62] until quite recently, when researchers recognized them as potential mechanisms for regulating multiple biological processes *in vivo*, such as transcription, replication, translation and telomere maintenance.

In this section, the structure, topology and biological roles of G-quadruplexes will be reviewed as well as their budding nanotechnological applications.

1.2.1. Structure and topology

Guanine-rich regions of nucleic acids (*G-tracts*) have a high propensity to self-associate into four-stranded helical structures known as *G-quadruplexes*. Their building blocks are planar aromatic arrays of four Hoogsteen hydrogen-bonded guanines, called *G-quartets* or *G-tetrads*, in which each guanine base forms two hydrogen bonds with its neighbours. Then, a stable G-quadruplex arises from the stacking of at least two of these G-quartets, which are held together by loops created by the nucleotides not involved in the tetrads.

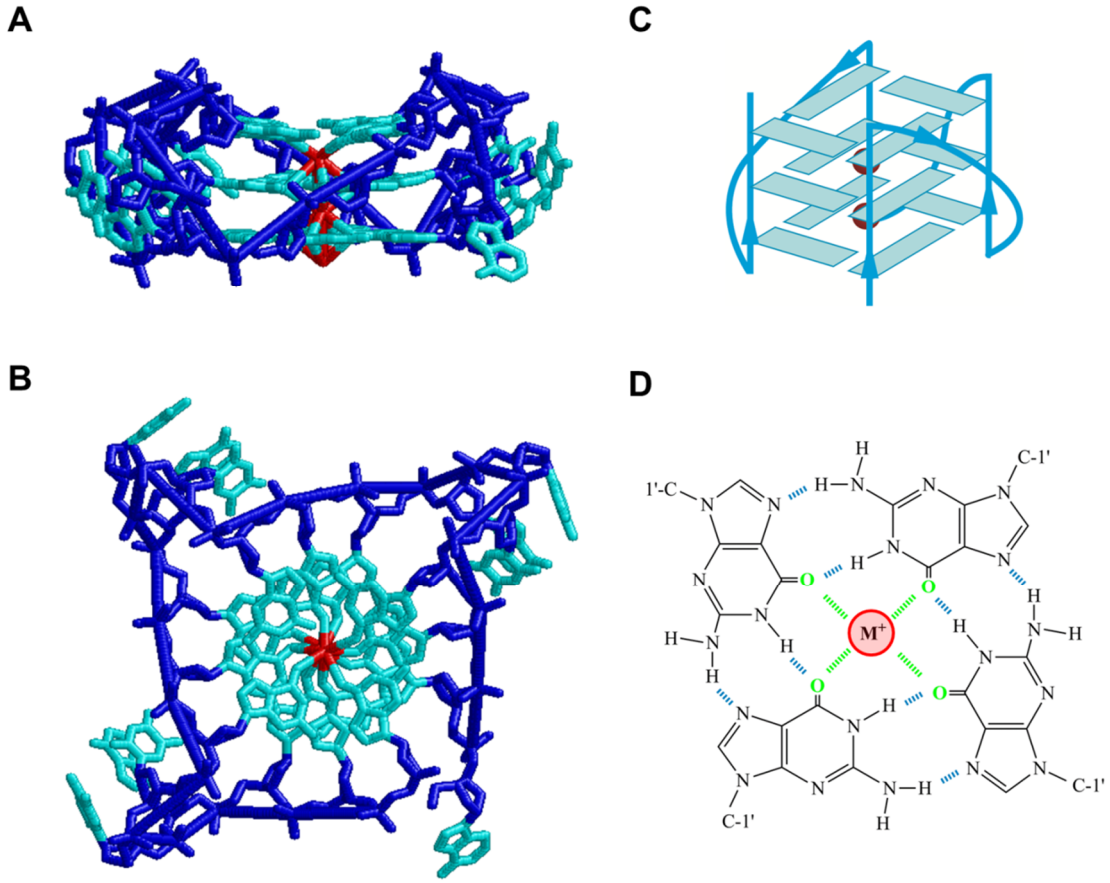

FIGURE 1.10. G-quadruplex structure. Front **(A)** and top **(B)** views of the crystal structure of the intramolecular G-quadruplex formed in the human telomeric DNA sequence $AGGG(TTAGGG)_3$ obtained from potassium-containing solution (PDB ID: 1KF1) [63]. Nitrogenous bases are represented in light blue, potassium ions in red and the backbone of alternating phosphate and pentose residues in dark blue. **(C)** Schematic plot of the same structure shown in (A), where blue rectangles represent guanine bases and red spheres, potassium ions. **(D)** Chemical structure of the basic structural motif of a G-quadruplex conformation, i.e. the G-quartet. Hydrogen bonds and the negatively charged channel created by O6 atoms (in green) are shown as blue and green dashed lines, respectively. M^+ stands for metal ion.

The stability of this stacking is strongly dependent on monovalent cations (M^+). It has been ascribed to the strong negative electrostatic potential created by the guanine O6 oxygen atoms, which results in a central negatively charged channel inside the G-tetrad stack where cations are located. The precise position of these counter-ions depends on their nature. For instance, potassium and ammonium ions tend to be situated equidistant from two consecutive tetrads so that they are eight coordinated to O6 atoms in a symmetric tetragonal bipyramidal configuration. The smaller sodium ion is able to being either in the plane of a G-quartet or accommodated midway between two successive ones [64-70].

In principle, G-quadruplexes can be folded from a single G-rich strand intramolecularly, being required four or more G-tracts in the strand, or by the intermolecular association of two (bimolecular) or four (tetramolecular) separate strands (*Fig.1.11*). More in detail, intramolecular G-quadruplex-forming sequences can be described as $G_mX_nG_mX_oG_mX_pG_m$, where *m* is the number of G residues in each short G-tract and X_m, X_o and X_p can be any combination of residues, including G, of equal —as in vertebrate telomeric sequences— or unequal length, that form the loops. Meanwhile, bimolecular quadruplexes reported to date respond to the association of two identical sequences $X_nG_mX_oG_mX_p$ —where n and p may be zero— whereas tetramolecular configurations would be formed by four $X_nG_mX_o$ or $G_mX_nG_m$ strands associating together [69, 70].

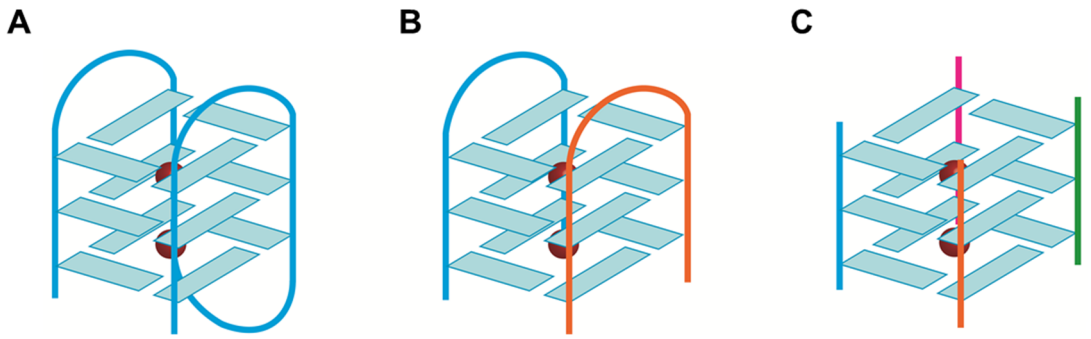

FIGURE 1.11. Types of G-quadruplexes according to the number of strands involved. G-quadruplexes can be folded from **(A)** a single G-rich sequence intramolecularly or by the intermolecular association of **(B)** two or **(C)** four separate strands (highlighted in different colours).

These non-canonical conformations exhibit extensive structural polymorphism relative to duplex DNA. The relative arrangement of strand polarity together with variations in loop size and sequence and/or in the number of stacked G-tetrads lead to a multitude of G-quadruplex structures, as indeed is found experimentally [63, 71-83].

Regarding the polarities of the four strands in a G-quadruplex, they can be parallel, three parallel and one anti-parallel, or alternating parallel, resulting in different conformations named as *parallel*, *3+1* or *hybrid* and *antiparallel*, respectively (see *Fig.1.12*). These variations in strand polarity affect the location of the loops between G-rich segments in intramolecular and bimolecular G-quadruplexes.

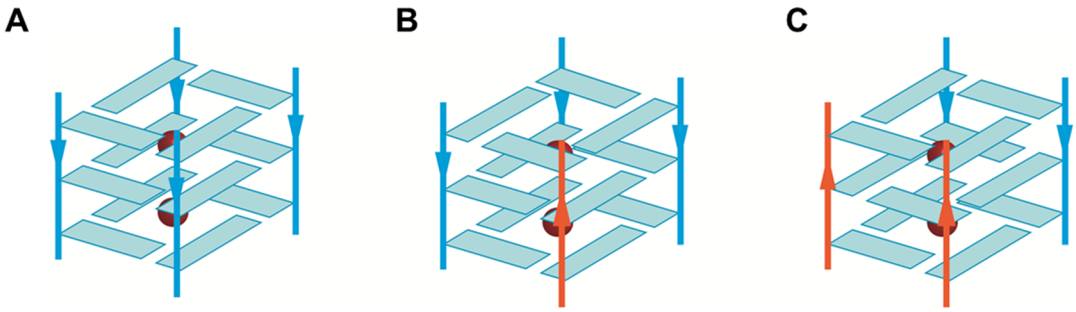

FIGURE 1.12. G-quadruplex folds depending on strand polarity: **(A)** Parallel; **(B)** 3+1 or hybrid and **(C)** antiparallel. Blue rectangles and red spheres represent guanine bases and metal ions, respectively. Strands highlighted in orange represent a change in polarity. For simplicity, *anti* and *syn* conformations of guanosine nucleosides are not taken into account in these schematic plots.

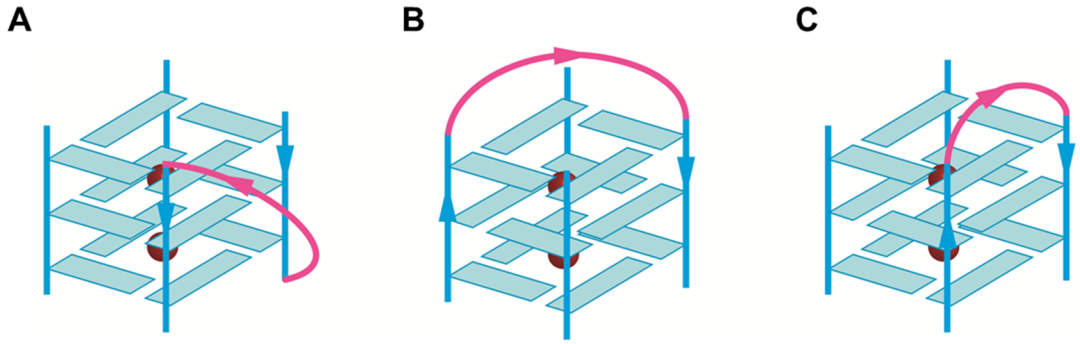

FIGURE 1.13. Types of loops in G-quadruplex structures. (A) Propeller-type; **(B)** diagonal-type and **(C)** lateral-type. For simplicity, *anti* and *syn* conformations of guanosine nucleosides are not taken into account in these schematic plots.

As shown in *Figure 1.13*, adjacent parallel strands require a connecting loop to link the bottom G-quartet with the top one, leading to *propeller* type loops. Anti-parallel G-strands can be linked either by *diagonal* or *lateral* (sometimes termed *edge-wise*) loops, depending on whether the strands are adjacent or diagonally opposed. In addition to its function as a linker, loops can themselves form stacking and hydrogen bonding interactions through their residues, further stabilizing (or not) G-quadruplex folds [65, 81].

In addition to the location of the loops, the orientation of individual strands in a G-quadruplex has also consequences for the glycosidic angle conformations —*anti* or *syn* (see *Fig.1.14*) —, of the guanosines involved in a G-tetrad. Thus, an all parallel orientation for all four strands, as in the tetramolecular G-quadruplex shown schematically in *Figure 1.12 (A)*, requires that all glycosidic angles adopt an *anti* conformation. Antiparallel G-quadruplexes have both *syn* and *anti* conformations, arranged in a way that is characteristic of a given topology and a set of strand orientations [69, 84].

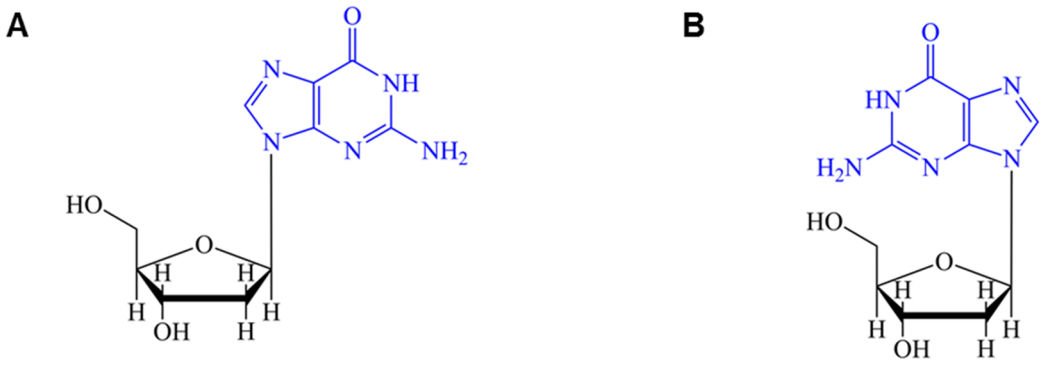

FIGURE 1.14. Glycosidic angle conformations in a G-tetrad. (A) Anti- and (B) syn- conformations of the guanosines involved in a G-tetrad.

1.2.2. Location of potential G-quadruplex-forming sequences (PG4s)

Both prokaryotic and eukaryotic genomes are rich in potential G-quadruplex-forming sequences (PG4s), even some human DNA and RNA viruses [85-93]. Thanks to genome-wide analyses, the location of these sequences has been revealed recently. Specifically, it was found that, in addition to being non-randomly distributed, pG4s are evolutionary conserved, indicating a selection pressure to retain these sequences at specific functional genomic domains [94].

Focusing on human genome, it contains almost four hundred thousand of PG4s, being found the highest concentration at telomeres. Inside the cell nucleus, other opportunities to form G-quadruplexes arises during DNA replication, transcription and repair, when DNA becomes transiently single stranded due to the breaking of Watson-Crick base pairings [95, 96]. RNA is also known to form these non-canonical structures, having been reported as many as 3000 mRNAs with PG4s located in their 5' untranslated regions (UTRs) [47-51]. Long G-rich RNA transcripts of telomeric DNA (TERRA) are also identified as genomic regions capable of folding into G-quadruplex structures [52].

1.2.2.1. G-quadruplexes in human telomeres

Telomeres are nucleoprotein structures located at the end of linear chromosomes that are essential to protect these regions from recombination and degradation activities [97]. In vertebrates, they consist of long tandem arrays of a short oligonucleotide sequence (~10-15 kb long) —TTAGGG in one strand (so-called *TG strand*) and AATCCC in the complementary [98] —, and an associated protein complex, thermed *shelterin* [99]. As shown in *Figure 1.15*, the TG strand of each telomere is longer than its complement, leaving a region of G-rich single stranded DNA (*G-overhang*) of up to a few hundred nucleotides (150-200 nt) at the 3' end. The 3'- OH group at the termini of this G-overhang is specifically recognised by *telomerase*, a reverse transcriptase in charge of telomere elongation in those cell types in which it is expressed [100, 101].

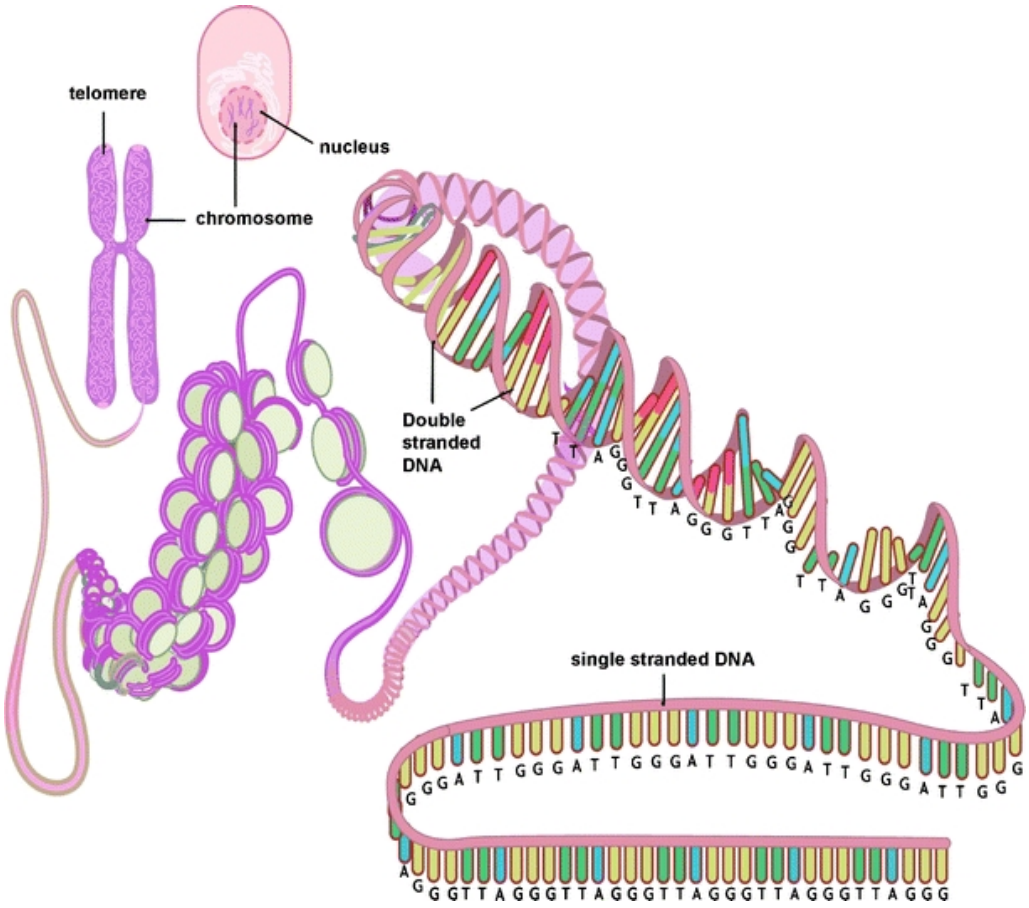

FIGURE 1.15. Telomeres at the end of linear chromosomes. Here, it is shown their structure in vertebrates: long tandem arrays of double stranded TTAGGG repeats (~10-15 kb long) in which the TG strand is longer than its complement, leaving a region of G-rich single stranded DNA overhang of up to a few hundred nucleotides (150-200 nt) at the 3' end. Reproduced with permission from [102].

Telomere shortening and telomerase activity are directly related to cellular aging and cancer [103-108]. After each cell division, telomeres shorten owing to the so-called "end-replication problem"; that is, the incomplete replication of linear chromosomes by conventional DNA polymerases. Most somatic cells do not express telomerase activity so when a critical telomere length is reached, they undergo a growth arrest called *replicative senescence*, leading to chromosomal instability and subsequent loss of cell viability. In contrast, high levels of telomerase have been detected in approximately 90% of human malignant tumour cells, enabling their division indefinitely [103]. Consequently, telomerase has become both a highly attractive biomarker and target for the development of new strategies for cancer detection, diagnosis and treatment [104, 108, 109].

Telomere synthesis by the telomerase requires that G-overhangs remains single-stranded. However, to protect themselves from nucleases and the enzymes that repair double-stranded breaks, each of these single-stranded regions can either fold back and pair with its complement in the duplex region of the telomere, forming what is known as a *t- loop* [97, 110, 111], or adopt a G-quadruplex conformation spontaneously [112-115]. These two capping structures need not be mutually exclusive and could represent different functional states [95].

In addition to the formation of DNA G-quadruplexes, there is also evidence for the presence of telomeric RNA G-quadruplexes in living cells [52]. Telomeres had always been thought to be transcriptionally silent until the recent finding that they are transcribed from the subtelomeric regions towards chromosome ends into **TElomeric Repeat-containing RNA (TERRA)** [116, 117]. These large, heterogeneous and noncoding transcripts contain subtelomere-derived sequences and an average of 34 GGGUUA repeats at their 3'end [118], which are susceptible to fold into multiple G-quadruplex blocks. Its biological importance lies in acting as a scaffold for the assembly of telomeric proteins involved in telomere maintenance and protection as well as in heterocromatin formation [119, 120].

1.2.3. Biological relevance

The recent evidence that G-quadruplex structures occurs *in vivo* suggests that these DNA and RNA higher order structures may play important cellular regulatory roles, such as protecting telomere ends against nucleases or specifying the initiation site of DNA replication. However, G-rich sequences also pose particular problems for cells, being necessary spatiotemporal regulation mechanisms for the formation and resolution of G-quadruplexes.

In telomeres, G-quadruplex structures provoke the inhibition of telomerase activity, leading to telomere shortening [121]. Then, a G-quadruplex folding and unfolding regulation mechanism is needed to allow telomere synthesis by the telomerase when required [122]. The phosphorylation of *telomere end-binding protein β* (TEBPβ) in ciliates or the preferential function of several RecQ helicases at vertebrate telomeres, such as *Werner syndrome* (WRN) and *Bloom syndrome* (BLM) proteins, would be some examples of that [112, 123-127].

Similarly, during replication, G-quadruplexes formed when DNA is transiently single-stranded need to be resolved by specialized DNA polymerases or unwinding helicases, like *Fanconi anemia complementation group J* (FANCJ) helicase, to allow replication [128, 129].

In transcription, the presence of a G-quadruplex could either impair its initiation by the RNA polymerase or inhibit the process if it is present in the antisense strand.

Finally, G-quadruplexes formed in the 5'-UTR of mRNAs can regulate translation and lead to aborted RNA transcripts in hexanucleotide repeat expansion diseases, such as amyotrophic lateral sclerosis (ALS) or frontotemporal dementia (FTD) [49, 130].

1.2.4. Applications

Since the demonstration of their presence in living cells [112, 114], G-quadruplex structures have raised a lot of interest in the field of nanomedicine. Thanks to their unique geometry, these non-canonical structures have become important drug-design targets for the treatment of various human disorders, ranging from viral infections to cancer [131, 132].

The specific recognition of G-quadruplexes by small ligands can occur through various binding modes in a manner analogous to that of double-helical DNA intercalators: *external stacking mode*, just on the surface of the terminal G-quartet (*Fig. 1.16 (A)*); *intercalating mode*, between the stacks of G-tetrads (*Fig. 1.16 (B)*) or *groove binding mode* (*Fig. 1.16 (C)*). Nonetheless, not all of them are considered to be equally probable. Intercalating binding of the ligand could be considered as the most improbable scenario due not only to the high stability and rigidity of the G-quadruplex structure, but also to the large amount of energy needed to distort their structural integrity. Thus, stacking of the drug on the outer planes of G-tetrads or groove binding seem to be more energetically favourable and probable [132]. Cationic porphyrin (TMPyP4) [133], acridine (BSU6039) [134], polycyclic acridine (RHPS4) [135] and several peptide-hemicyanine conjugates [136] are some examples of the compounds that have

been identified as interacting specifically with DNA G-quadruplexes following the latter patterns.

Over the last decade, studies about stabilization of G-quadruplex structures by small ligands have been mainly focused in the G-rich single-stranded regions of human telomeres. As described previously, each 3'-telomeric overhang is capable of adopting a G-quadruplex conformation, blocking the elongation step catalysed by telomerase and, consequently, leading to cellular senescence. Then, achieving stable G-quadruplex configurations in these regions by using small ligands has become a compelling antitumoral strategy [137-146]. But, without any doubt, the major challenge will be identifying nontoxic drugs that are truly selective for G-quadruplex assemblies [146].

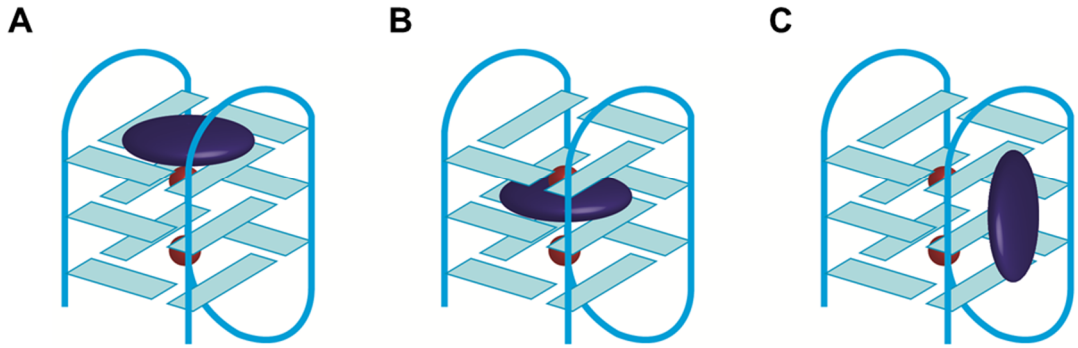

FIGURE 1.16. Interaction modes between ligands and G-quadruplexes: **(A)** external stacking mode, on the surface of the terminal G-quartet; **(B)** intercalating mode, between the stacks of G-tetrads, and **(C)** groove binding mode. Red spheres represent metal ions and dark blue disks, GQ-binding drugs.

Besides their potential application as antitumoral agents, G-quadruplexes are also considered promising candidates in the field of nanotechnology. In particular, G-quartets have been proposed as building blocks of new electronic nanodevices due to their capability to form higher order structures, termed G-wires, as well as their rather high conductance [147-154]. Furthermore, some G-quartet-forming DNA aptamers have been reported as attractive stationary phase materials in bioanalytical chemistry [155-157], effective catalysts [158, 159] and, labelled with fluorescent dyes, as new types of biosensors [160-167].

1.3. Mechano-chemistry and single-molecule manipulation techniques

The essential role that forces play in biochemical processes soon attracted the interest of many scientists, giving rise to a new, dynamic, exciting and interdisciplinary research field

called ***mechano-chemistry*** [168]. It includes the study of biochemical processes as diverse as protein folding [169-176], the protein-induced bending of DNA [177, 178], the stress-induced enzymatic catalysis [179, 180], the behaviour of molecular motors [181-187] and the elasticity of nucleic acid structures [45, 188-192].

The cell environment where all these fundamental processes take place is stochastic. It means that proteins and enzymes exist in a background of random thermal fluctuations that affect their dynamics. The study of these fluctuations in real time and at a single-molecule level was not possible until quite recently. Traditional techniques, such as circular dichroism (CD), nuclear magnetic resonance (NMR) and fluoresce microscopy, can only investigate chemical processes on a bulk level, which only gives access to the average dynamics. New experimental tools should be developed and, over the last two decades, outstanding progresses have been made.

In fact, ***single-molecule manipulation techniques*** **(SMTs)** have emerged as powerful tools to investigate the movements, forces and strains generated by single biological molecules in real time [168, 193-197]. Their most important feature is the absence of problems associated with populating averaging. Whereas bulk experiments hinders the access to fluctuations and the identification of multispecies distributions, single-molecule approaches allow the access to this information as there are not constrictions associated to the multiple interactions that take place in an ensemble of equivalent molecules. Nowadays, a wide variety of SMTs can be distinguished depending mainly on force and displacement ranges and spatial and temporal resolution. The most commonly used are optical tweezers (OT), magnetic tweezers (MT) and atomic force microscopy (AFM). A comparison of them can be found on *Table 1.1*.

	Optical tweezers	**Magnetic tweezers**	**AFM**
Force range (pN)	0.1-100	0.01-10	$10\text{-}10^5$
Displacement range (nm)	$0.1\text{-}10^5$	$5\text{-}10^4$	$0.5\text{-}10^4$
Spatial resolution (nm)	0.1-2	5-10	0.5-1
Temporal resolution (s)	10^{-4}	$10^{-1}\text{-}10^{-2}$	10^{-3}
Practical advantages	Specific manipulation High force resolution	Specificity to magnets Ability to induce torque	High spatial resolution

TABLE 1.1. Comparison of the most commonly used single-molecule manipulation techniques. Based on Ref. [168] and [197].

In this section the operating principle of the SMTs used in this thesis, optical tweezers and

AFM, will be described in detail, being also highlighted their practical limitations and drawbacks.

1.3.1. Optical tweezers

Optical tweezers (OT) are extremely well suited for the measurement of force and motion in a large variety of biomolecular and cellular processes. It can be used to exert small forces on particles ranging in size from ~20 nm to several micrometres and, simultaneously, measure the three-dimensional displacement of the trapped particle with sub-nanometre accuracy and sub-millisecond time resolution. Single cells, organelles within cells, lipid vesicles and polystyrene or silica microspheres —used alone or as probes linked to a biomolecule of interest— are examples of trapped objects by an optical field [197-201].

1.3.1.1. Optical trap

The force sensor of this single-molecule technique is an optical trap. It is created by focusing a laser beam to a diffraction-limited spot with a high numerical aperture (NA) microscope objective [202]. Not all types of particles can be optically trapped; it is dependent on their refractive index. Whereas a dielectric particle is attracted to regions of higher field intensity, a metallic one may have either an attractive or a repulsive nature towards them [203]. Cells, organelles and biomolecules have been shown to behave as dielectric [204-207].

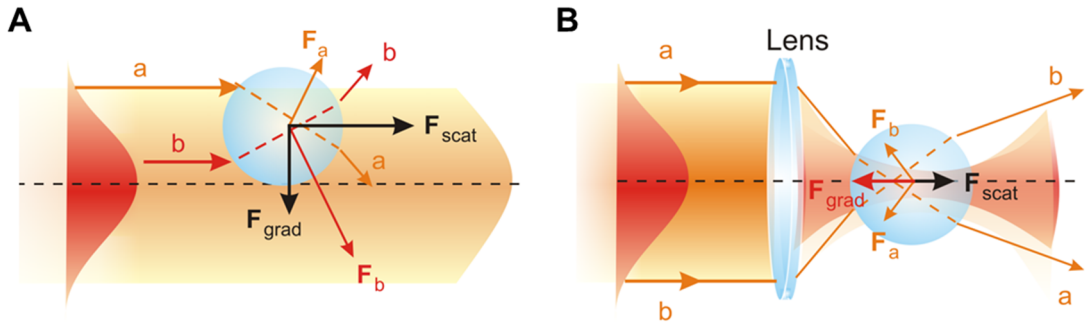

FIGURE 1.17. Optical forces acting on a dielectric particle. (A) Schematic representation of the scattering (F_{scat}) and gradient (F_{grad}) components of the optical force on a dielectric sphere originated by a Gaussian laser beam. **(B)** Axial stability of a dielectric sphere on a focused laser beam due to the gradient force along the centre line of the beam, towards the focal region. Reproduced with permission from [208].

Two kind of forces act on an optically trapped dielectric particle: the *scattering force*, produced by the photons striking the particle along their propagation direction, and the *gradient force*, produced by a gradient of field intensity. To form a stable optical trap, the gradient force

along the optical axis must overcome the scattering force, which is achieved if the focusing of the laser beam is strong enough (see *Fig.1.17*).

For small displacements (~ 150 nm) of the trapped object from its equilibrium position, the force is linearly proportional to the displacement, so the optical trap can be interpreted as a Hookean spring. The spring constant, or stiffness, depends on how tightly the laser is focused, the laser power and the polarizability of the trapped object and it needs to be well characterized to attain accurate force data [197, 209].

1.3.1.2. Optical tweezers set-up: Basic elements and up-to-date configurations

Although the range of OT experimental configurations is diverse and keeps growing, there are some basic elements that are present in all of them:

- **Laser illumination**. For high-precision optical trapping measurements, a laser beam with a Gaussian output intensity profile and superior power stability is required. Thus, the achievement of the smallest focal spot is possible, producing the largest and most stable optical gradient and reducing drastically measurement noise. Moreover, to minimize optically induced damage in biological samples, a trapping laser operating in the nearly-infrared region from 800 to 1100 nm is typically selected. In that optical window, water and biological specimens exhibit an adequate transparency. Dual-beam systems, which basically consist of two equal counter-propagating laser beams brought to the same focus, are also proposed as an alternative to reduce photodamage [210]. Regarding the output power, a couple of hundred of mW is generally appropriate to generate forces up to hundreds of pN. [197, 209]

- **Strongly-focusing objective lens**. To achieve the steep focus needed to create a stable optical trap, the NA of the trapping microscope objective should be near 1.2. This requirement is met by both water and oil immersion objectives. However, and although the latter offers the highest NA, water immersion objectives are the most commonly used for optical trapping since they do not introduce spherical aberrations that degrade optical trap performance deep in solution. [197, 211, 212]

- **Microfluidics chamber**. It is the place where the single-molecule experiments are carried out in buffer solution at a controlled rate. It is usually built by using two glass coverslips sandwiching two parafilm layers with an imprinted system of channels (see

Fig.1.18). Single cells or coated beads with chemically attached biomolecules are introduced inside the fluidic chamber by the upper channel and flowed to the focal region, located in the central one, through a glass capillary. In some OT configurations, a micropipette is placed inside the chamber not only to act as the frame of reference of the optical trap, but also to hold a second coated bead by suction. This second microsphere, to which the biological specimen will also bind chemically, is flowed through the lower channel, also connected with the central one by a glass capillary.

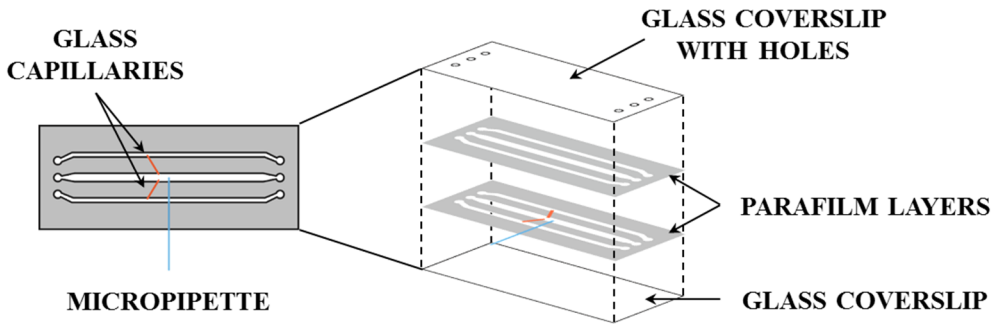

FIGURE 1.18. Schematic representation of a microfluidics chamber. Microfluidics chambers are usually built by using two microscope coverglasses sandwiching parafilm layers with imprinted channels.

- **Condenser lens**. It is used to collect the light scattered by the optically trapped biological specimen and convert it into a parallel beam.

- **Light-detection system**. The interference between the light scattered by the trapped bead and unscattered light is measured with a quadrant photodiode or position-sensitive photodetector (PSD) to further obtain the force and displacement measurements that experience the sample of interest. When a trapped bead is at its equilibrium position, a symmetric interference pattern and a null detector signal is produced. Any displacement of the bead results, then, in an asymmetric interference profile, which generates a detector signal proportional to the bead displacement. Therefore, registering the light deflected by the bead and characterizing the optical trap as a spring constant, both bead movement and force can be determined. [197]

A schematic representation of the OT set-up previously described and used in this thesis (see *Chapter 3*) is shown in *Figure 1.19.*

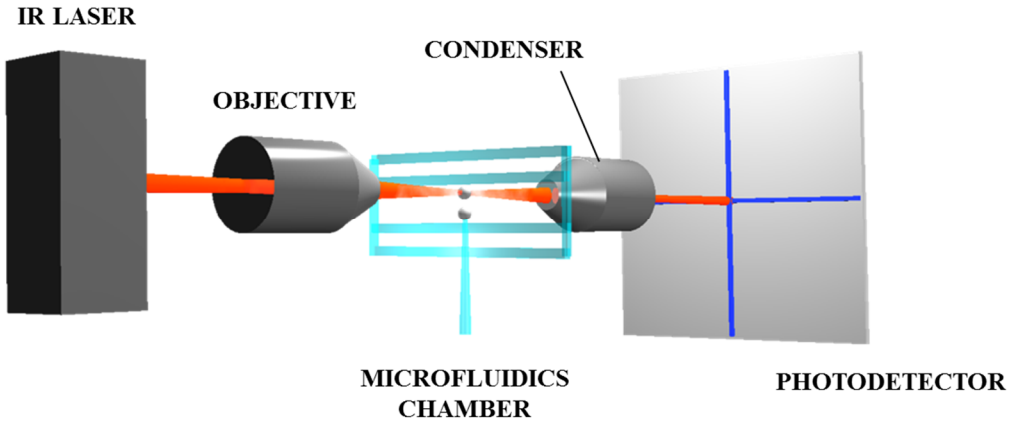

FIGURE 1.19. Basic elements in a OT instrument. Laser light (typically infrared) is highly focused by a microscope objective creating an optical trap inside a microfluidic chamber, where the single-molecule experiments are carried out. A bead or a single cell can be trapped in the focal region whereas a second bead is held by suction on top a micropipette. The light scattered by the sample in the trap is firstly collected by a condenser lens and, then, redirected to a position-sensitive photodetector.

Nowadays, more sophisticated OT instruments are pursued by several laboratories around the world, looking for better and high-resolution optical trapping measurements. The so-called miniTweezers set-up [213], also used in this thesis (*see Chapter 3*), is an example of that, which has been subsequently improved by adding a temperature controller [214]. Within this new generation of OT systems are the hybrid configurations; i.e., the combination of OT with other SMTs to take advantage of the best features of each technique. Magneto-optical tweezers [215] —which combines the nanometer resolution of OT and the easy manipulation of MT— and OT-FIONA (*Fluorescence Imaging with One Nanometer Accuracy*) instrument [216] —which allows the detection of biomolecule position, conformation or biochemical state simultaneously with its mechanical manipulation— are particularly fruitful examples of these hybrid tools.

1.3.1.3. Single-molecule manipulation of nucleic acids

Thanks to the advent of the OT technique, elastic properties and folding dynamics of diverse nucleic acid structures have been investigated from their characteristic force-extension curves (*F-x curves*): from double-stranded DNA and RNA polymers at different environmental conditions [28, 29, 45, 217, 218] to non-canonical conformations, like G-quadruplexes [219-223].

These single-molecule experiments are usually carried out in an experimental OT set-up like that shown in *Figure 1.19*. A single molecule is grabbed by two different, specific

sites by means of two short nucleic acid stretches, usually called handles, to properly functionalized dielectric microspheres. Typically, nucleic acid fibres are prepared containing modified nucleotides with biotin groups at one end and digoxigenin at the opposite, so that they can bind to stretptavidin- and anti-digoxigenin-coated beads, respectively (see *Fig.1.20*). The purpose of the handles is to avoid direct contact between the molecule of interest and the surfaces of the beads, which may produce artefacts owing to surface effects [224]. F-x curves are obtained by the stretching, and subsequent relaxation, of a single biomolecule by moving the micropipette relative to the laser trap at a constant rate. The extension of the nucleic acid strain is determined from the distance between the beads with a video camera whereas the force acting on it is calculated from the displacement of the laser beams on the PSDs [218].

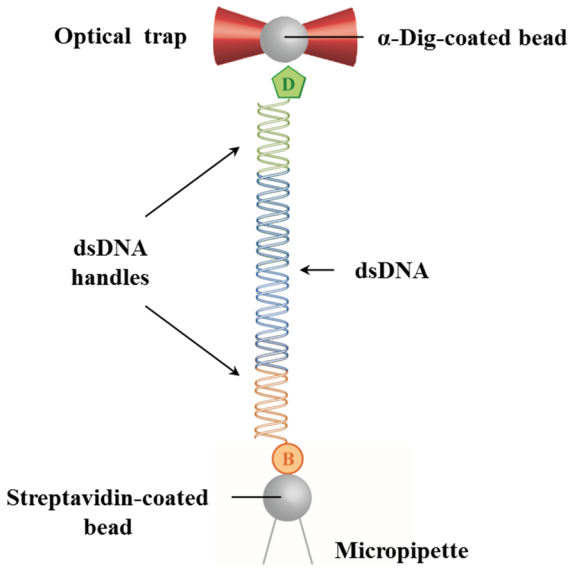

Figure 1.20. Scheme of the experimental configuration in OT measurements. A single dsDNA molecule was tethered by opposite DNA duplex ends between two polystyrene microspheres: an α-digoxigenin-coated bead, optically trapped, and a streptavidin-coated bead, held by suction on top of a micropipette. "B" stands for biotin and "D" for digoxigenin.

In *Figure 1.21* a typical F-x curve of a single double-stranded DNA molecule is shown. Three different elasticity regimes can be distinguished during its stretching pathway [30, 218]:

- ***Entropic elasticity regime***. Below some 5 pN, the molecule, initially bent over itself in a random way as a result of thermal fluctuations, is stretched like a Hookean spring until its characteristic contour-length; i.e. its length at maximum physically possible extension.

- ***Enthalpic or intrinsic elasticity regime***. From 5 to 60 pN, approximately, the molecule is stretched beyond its contour length (L_0).

- ***Overstretching transition***. At about 65 pN —in non-torsionally constraint molecules and depending on ionic strength, pH and sequence—, the molecule undergoes a structural transition to an almost totally unwound state with an elongation of 1.7 times its contour length. This sudden change happens over a very narrow force range (~2 pN) and involves an increase in the rise per phosphate from 3.4 to 5.8 Å. The bases unstack and, depending on environmental conditions, the strands may either remain linked by partial base pairing —like a parallel ladder (*S-form*)—, or separate into a totally molten state (*force-induced melting*) or in a heterogeneous combination of both [225-235]. The stretching cycle is always reversed before 100 pN because the biotin-streptavidin bond often breaks after a few seconds at this force.

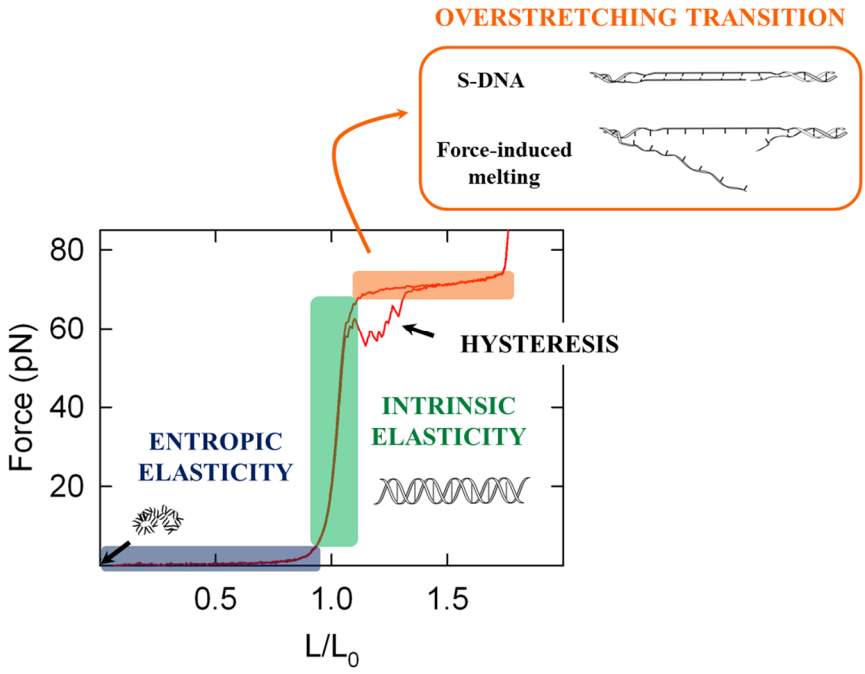

FIGURE 1.21. Characteristic force-extension curve of a single dsDNA molecule. The graph shows an individual dsDNA molecule from the [30286, 34286] fragment of λ DNA (47.3% GC, 4.0 kbp) in 150 mM NaCl, 150 mM NaCl, 10 mM Tris-Cl, 1 mM EDTA, pH 7.5. Insets illustrate the strand arrangement in each elasticity regime according to the literature (see text for details). Extension axis has been normalized to the contour length of the molecule (L_0).

When the stress on the molecule decreases, it begins to return to its natural state but following a slightly different pathway, provoking a certain hysteresis between the stretching and the relaxation traces. This absence of overlapping in the F-x curve is a hallmark of the

reannealing of the strands out of equilibrium. The DNA regions which have remained base-paired rapidly contract to the B-form but the frayed tracts take more time to reanneal, perhaps due to the secondary structure they could adopt in the overstretched molecule [218].

Regarding non-canonical nucleic acid structures, as G-quadruplexes, F-x curves provide a great deal of information, including their mechanical stability ($F_{rupture}$) and unfolding extension (Δx). Both parameters are extracted from the inherent rupture events, which become observable as jumps over the characteristic F-x curves of the handles —dsDNA, dsRNA or hybrids— (see *Chapter 3*).

1.3.1.4. Limitations and drawbacks

The awesome versatility and outstanding precision of this single-molecule technique are accompanied unfortunately by important limitations and drawbacks.

Some of these important difficulties are associated with the fact of using light to generate force. As trap stiffness depends on the gradient of the optical field, any optical perturbation in the intensity distribution affects the performance of the OT. As a result, high-resolution measurements are only achieved if trapped samples are optically homogenous.

OT also lack selectivity and exclusivity. Any dielectric particle near the focus region can be optically captured, being possible the simultaneous trapping of a quite large number of objects. For this reason, freely diffusing samples must be kept at extremely low concentrations to prevent this undesirable effect [197].

Local heating in the vicinity of the optical trap due to the high light intensity at the focus of the trapping laser (10^9-10^{12} $W \cdot cm^{-1}$) is another drawback but it can be quantified, both theoretically and experimentally [236-238], and even kept limited [214].

Optical damage induced in trapped biological specimens is less well understood. Although the use of trapping lasers operating in the near-infrared spectrum minimizes photodamage, singlet oxygen or other reactive oxygen species could persist throughout this wavelength range [198].

Apart from all these technical concerns, several practical matters also must be taken into account when using OT. As mentioned in *Table 1.1*, the range of applied forces with this technique is 0.1-100 pN. The lower limit is set by the lowest stiffness needed to ensure a stable optical trap whereas the upper one corresponds to the maximum power in the specimen plane. Regarding the range of motion that can be measured with a fixed optical trap, it is limited to ~ 400 nm or less, but the range over which the stiffness is constant and the optical trap can be characterized as a Hookean spring is considerably smaller (~150 nm). Last but not least, the artificial nature of the single-molecule linkage strategy has to be considered. Whereas individual cells can be held directly in an optical trap, handling single molecules requires the use of molecular handles as explained in the previous section. The absence of this type of configuration in living conditions is considered as a handicap for *in vivo* applications of this technique. [197, 239]

1.3.2. Atomic force microscopy

Atomic force microscopy (AFM) is a powerful and widely used imaging technique for visualizing the surface of samples with sub-nanometre resolution. Initially developed to overcome limitations of scanning tunnelling microscopy (STM) in imaging nonconductive samples [240, 241], AFM soon became an ideal tool for biological applications [242]. The possibility of operating under near-physiological conditions and manipulating individual molecules represent the major steps forward [243, 244].

1.3.2.1. Operating principle

Basically, an atomic force microscope consists of a soft cantilever of 100 to 200 μm in length with a sharp tip at its end —a pair of microns long and less than 100 Å in diameter—, which is used either to scan topography or to extend molecules attached to a surface [245].

When AFM is used in an **imaging mode** (*Fig. 1.22*), the cantilever scans the surface of the sample line after line. When the tip is brought into proximity of the surface, forces between the tip and the sample lead to a deflection of the cantilever according to Hooke's Law:

$$F = -kz \qquad [1.1]$$

where F is the force exerted on the cantilever, k the spring constant determined by the type of cantilever, and z the displacement experienced by the cantilever along a scanning

direction perpendicular to the surface of the sample. Typically, this deflection is measured using a laser spot reflected from the top surface of the cantilever into an array of photodiodes. This position-sensitive photodetector (PSD) allows the separation of the signals originated by the vertical and torsional movements of the cantilever during the scanning of a surface. The information of the topography of the sample resides in this vertical displacement whereas the frictional forces between the tip and the sample are provided by torsional data [246].

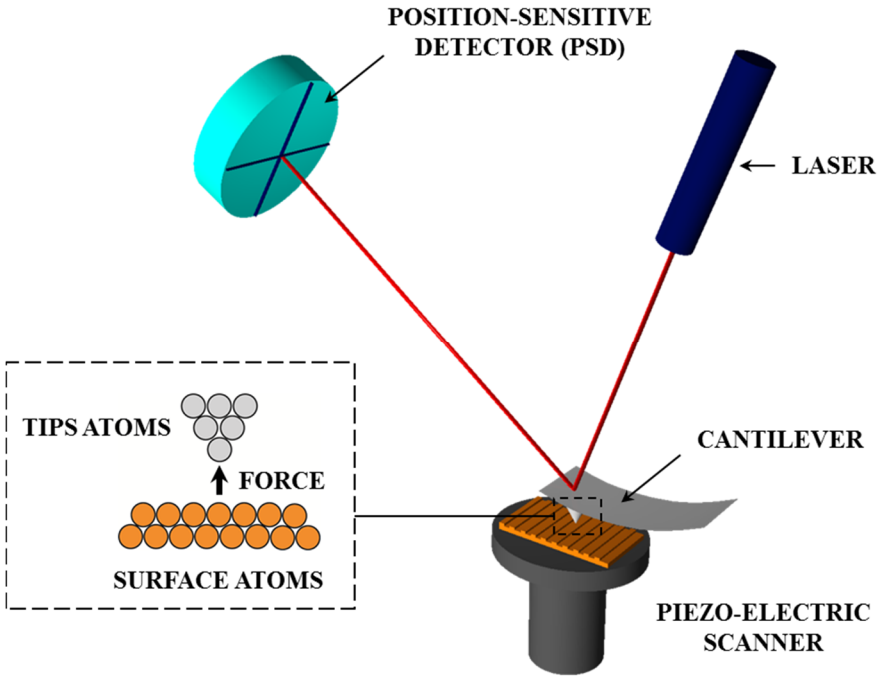

Figure 1.22. Basic AFM configuration operating in imaging mode. The force originated by the interaction between the atoms of the tip and those of the surface is a signal strongly dependent on the distance, which is recorded by a photodetector through the deflection of the cantilever during scanning.

When AFM is used for manipulating single molecules (*Fig.1.23*), this specialized version is called a ***molecular force probe*** (**MFP**) and its use is fundamentally distinct from imaging AFM. Specifically, the cantilever tip is moved only vertically with respect to the surface by piezoelectric actuators. After the initial contact with the surface, the cantilever is pressed into the surface on which the sample is deposited with a predetermined constant force. The tip is then retracted and, because of the attachment of the sample (DNA, RNA or proteins) to it, the cantilever bends toward the surface. The displacement of the cantilever is monitored directly with a PSD. It is important to note that both the cantilever and the surface have to be chemically modified to achieving these specific attachments to the ends

of the biomolecule of interest.

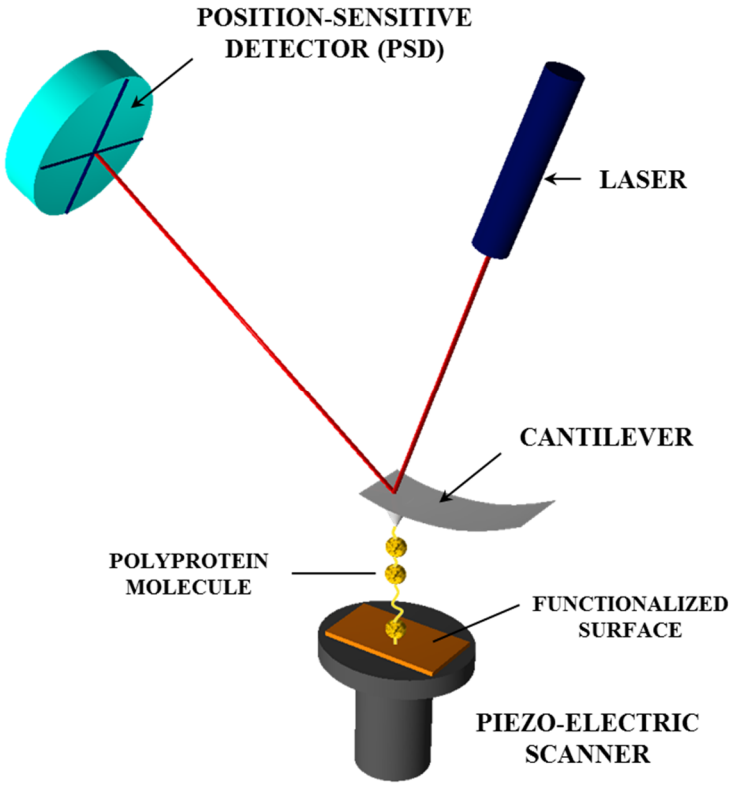

Figure 1.23. Basic AFM configuration operating as a molecular force probe. A typical AFM pulling experiment is represented in which a polyprotein molecule (displayed in yellow) is attached to a sample surface (displayed in orange) and an AFM cantilever tip. The piezoelectric scanner is retracted along the axial direction, increasing the separation between the cantilever and the sample surface. The force on the molecule is provided by the cantilever deflection, and the extension of the molecule is equal to the separation between the AFM tip and the sample surface.

As a result of MFP measurements, two types of characteristic data can be obtained: force and extension. The force exerted on the molecule during its pulling can be calculated by Hooke's law, using as z the cantilever deflection, and its extension is equal to the separation between the MFP tip and the sample surface. The precision of force and extension data depends on the quality of the calibration of the cantilever and the piezoelectric actuator, respectively. Piezoelectric scanners used in MFPs provide angstrom-level resolution and, together with a closed-loop position feedback control, are accurate and not susceptible to drift and hysteresis.

1.3.2.2. Limitations and drawbacks

The main drawbacks and limitations of AFM stem from the large size (100-200 μm in length) and relatively high stiffness (10- 10^5 pN/nm) of the cantilevers. Consequently, a large minimal force is imposed, hindering AFM applications in the study of many biological processes and structures in which the forces involved are so minute.

In addition, unlike OT [207, 239], AFM is inappropriate for experiments *in vivo* because the AFM tip cannot penetrate the cell membrane without breaking it.

Chapter 2

TERRA G-quadruplexes imaged by atomic force microscopy

2.1. Introduction

Numerous studies have proven the utility of atomic force microscopy (AFM) for imaging and analysing nucleic acid structures at the single-molecule level, even under near-physiological conditions [45, 148, 247-261]. The data collected can be displayed as a 3-D topography of the sample, allowing sub-Angstrom accuracy in the determination of the z dimension of the sample of interest, i.e. height. Here, AFM has been used, for the first time, to make direct measurements of human telomeric repeat-containing RNA (TERRA) molecules in ambient air.

TERRA molecules are involved in various cellular regulatory functions, such as regulation of telomere length [117, 262, 263], inhibition of telomerase [264], telomeric heterochromatin formation [119, 265] and telomere protection [116, 266-268]. As their DNA counterparts, they are able to fold into multiple G-quadruplex (GQ) blocks, each formed by a four repeat segment [52, 258, 269-271]. Until now, several AFM studies have been done on the possible arrangements of GQs along human telomeric DNA [256, 258]. However, direct visualization of TERRA GQs formed in long RNA molecules by AFM had not been reported yet. We have designed different RNA substrates with telomeric GQ-forming sequences, proposing an efficient method to check the formation of these special structures before further analysis.

2.2. Materials and methods

2.2.1. Synthesis of TERRA molecules

Three different RNA substrates with telomeric GQ-forming sequences were visualized by AFM. In particular, two of them —*TERRA-25 and TERRA-25 with double stranded RNA*

handles— contain twenty five GGGUUA repeats, giving rise to a maximum of six GQs, whereas the other one —*TERRA-16*— only has sixteen, being able to form four GQs at the most. The synthesis protocol of each of them is detailed below.

2.2.1.1. TERRA-25 molecules

The DNA template for this TERRA substrate was obtained by PCR (*Polymerase Chain Reaction*) amplification of a conveniently modified pUC18 plasmid (with an insert of twenty five GGGTTA repetitions) using the following pair of primers:

F1: 5'-Sp6 promoter-GAATAAGGGCGACACGGAAATGTTG-3'
R1: 5'-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one. This technique, developed by Kary Mullis [272], is widely used in molecular biology to amplify a specific DNA segment in an easy, cheap and reliable way, generating thousands to millions of copies in a few hours and without using living cells. The reaction mixture (50 µl) and PCR cycling conditions used here are detailed in *Table 2.1* and *Table 2.2*, respectively.

Reagent	V(µl)
Ultrapure Milli-Q-filtered water	22.0
Modified pUC18 plasmid (10 ng/µl)	1.0
Primer F1 (Integrated DNA Technologies, 10 µM)	1.0
Primer R1 (Integrated DNA Technologies, 10 µM)	1.0
SapphireAmp® Fast PCR Master Mix (Takara ClonTech, Ref. RR350A)	25.0

TABLE 2.1. PCR mixture for DNA telomeric GQ-forming sequence amplification.

Cycle		Temperature (ºC)	Time (min)
Hot Start Automatic		94.0	3
Start cycle (x 40)	Denaturation	95.0	0.5
	Annealing	55.0	0.5
	Elongation	72.0	1.5
Temperature		72.0	5
Store		4.0	∞

TABLE 2.2. PCR cycles for DNA telomeric GQ-forming sequence amplification.

The obtained reaction product was purified using *Wizard® SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting the sample in 50 µl of nuclease-free water. Subsequently, it was in vitro transcribed by Sp6 RNA polymerase (New England Biolabs, NEB,

Ref. M0207S; 3h, 40ºC), whose promoter sequence was incorporated at the 5' end of the forward PCR primer (F1). The generated transcript was purified using *RNeasy MinElute Cleanup Kit* (Qiagen, Ref. 74204), eluting the sample in 20 µl of RNase-free water. Then, the sample was equilibrated in 15 mM KPi pH=7 buffer by diafiltration using *Amicon® Ultra centrifugal filter units* (Merck Millipore, Ref. UFC901024). Finally, to favour the formation of the G-quadruplexes, the sample is subjected to heating at 90ºC for 5 min, after which it is gradually cooled at room temperature.

Reagent		**V(µl)**
RNase-free water		14.80
Sp6 polymerase buffer 10X (NEB, Ref. M0207S)		2.50
rNTPs dil 1:10	A	1.25
	C	
	G	
	U	
DNA template (~50-60 ng/µl)		1.20
RNase Inhibitor (NEB, Ref. M0314S)		0.50
Sp6 polymerase (NEB, Ref. M0207S)		1.0

TABLE 2.3. Mixture reaction for in vitro transcription of DNA telomeric GQ-forming sequences.

A schematic representation of the whole process is shown in *Figure 2.1*.

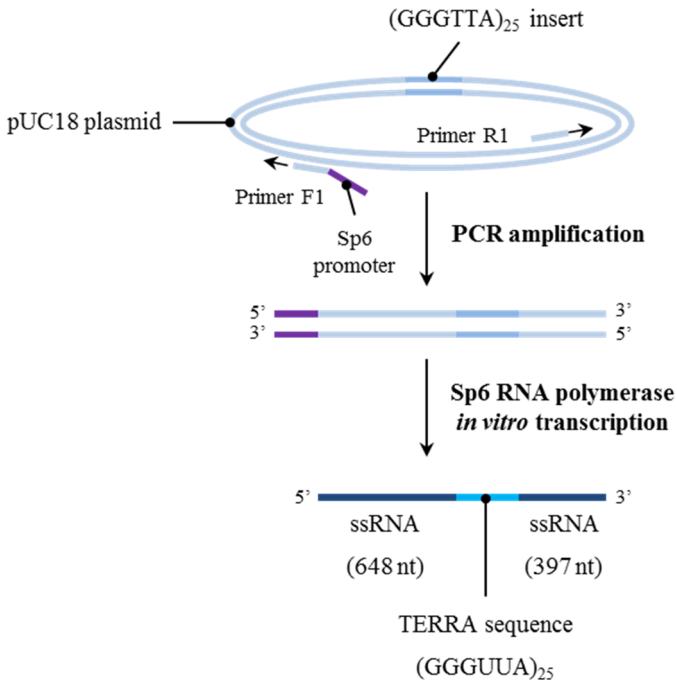

FIGURE 2.1. Schematic representation of the steps involved in the synthesis of TERRA-25 molecules. DNA template is generated by PCR amplification of a conveniently modified pUC18 plasmid and subsequently transcribed by Sp6 RNA polymerase, generating single-stranded (ss) RNA molecules with the desired number of telomeric repetitions.

2.2.1.2. TERRA-25 molecules with double-stranded RNA handles

To control that the structures visualized by AFM actually corresponded to GQs, TERRA-25 molecules were modified so that the telomeric sequence was flanked by asymmetric double-stranded RNA (dsRNA) at both sides. To synthesise the DNA templates for these handles, two more PCR amplifications were needed using the following pairs of primers:

F2: 5'-GAATAAGGGCGACACGGAAATGTTG-3'
R2: 5'-Sp6 promoter-GAGTCGTATTAAGCTTGGCACTGG-3'
F3: 5'- CTTCGAATTCGTAATCATGGTCATAGCTGT-3'
R3: 5'- Sp6 promoter-GGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the numbers 2 and 3 refer to handle A and handle B, respectively. Reaction mixtures and PCR cycling conditions are the same as previously described in *Table 2.1*, *2.2* and *2.3*.

The final assembly is achieved mixing equimolar amounts of the three transcripts (main molecule, handle A and handle B) and annealing them in the presence of 100 mM potassium chloride, which favours the GQ stabilization, in a PCR apparatus following the procedure specified in *Table 2.4*. A schematic representation of the whole process is shown in *Figure 2.2*.

Cycle	Temperature (ºC)	Time (min)
Denaturation	90.0	5
Annealing	85.0	10
	62.0	90
	52.0	90
	20.0	∞

TABLE 2.4. Annealing procedure for TERRA assemblies.

2.2.1.3. TERRA-16 molecules

To study the deposition geometry of the G-tandems with respect to the mica surface, TERRA molecules with fewer telomeric repeats were synthesised. Specifically, they contain sixteen GGGUUA repetitions; that is, molecules that can give rise to a maximum of four GQs. In addition, to favour their identification during AFM imaging, the 5' end of these single-stranded molecules was hybridized with a complementary RNA fibre to form a double helix following the same protocol described in the previous section. Then, two DNA templates were obtained by PCR amplification of a conveniently modified pUC18 plasmid (with an insert of

sixteen GGGTTA repetitions) using the former pairs of primers F1-R1, for the main single-stranded molecule, and F2-R2, for its 5'-end-complementary. Reaction mixtures and conditions for PCR, in vitro transcription and annealing procedures are the same as previously detailed in *subsection 2.2.1.2*. A schematic representation of the whole process is shown in *Figure 2.3*.

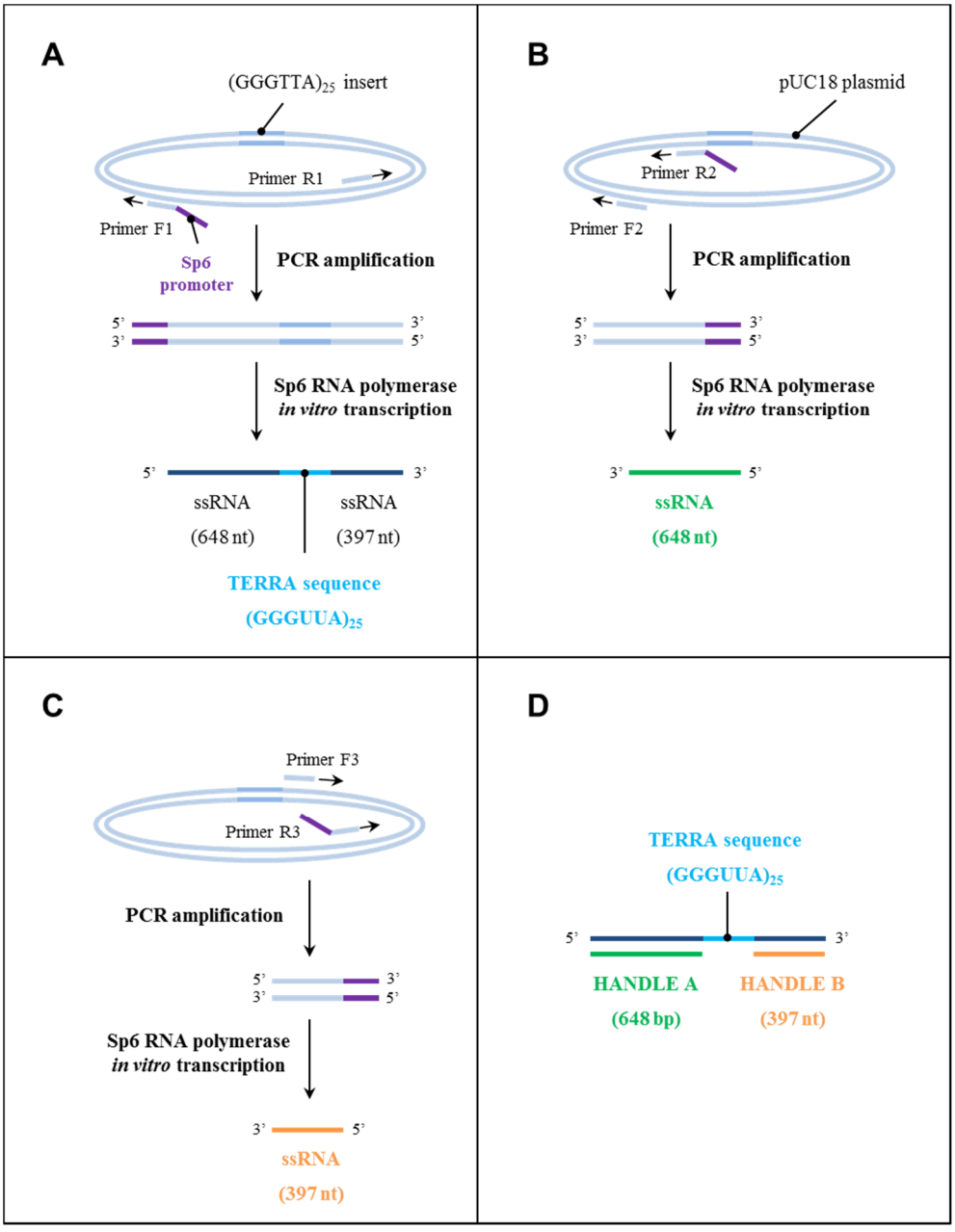

FIGURE 2.2. Schematic representation of the steps involved in the synthesis of TERRA-25 molecules with dsRNA handles. **(A)** Synthesis of TERRA sequence; **(B)** synthesis of ssRNA to construct handle A; **(C)** synthesis of ssRNA to create handle B and **(D)** annealing of the three transcripts.

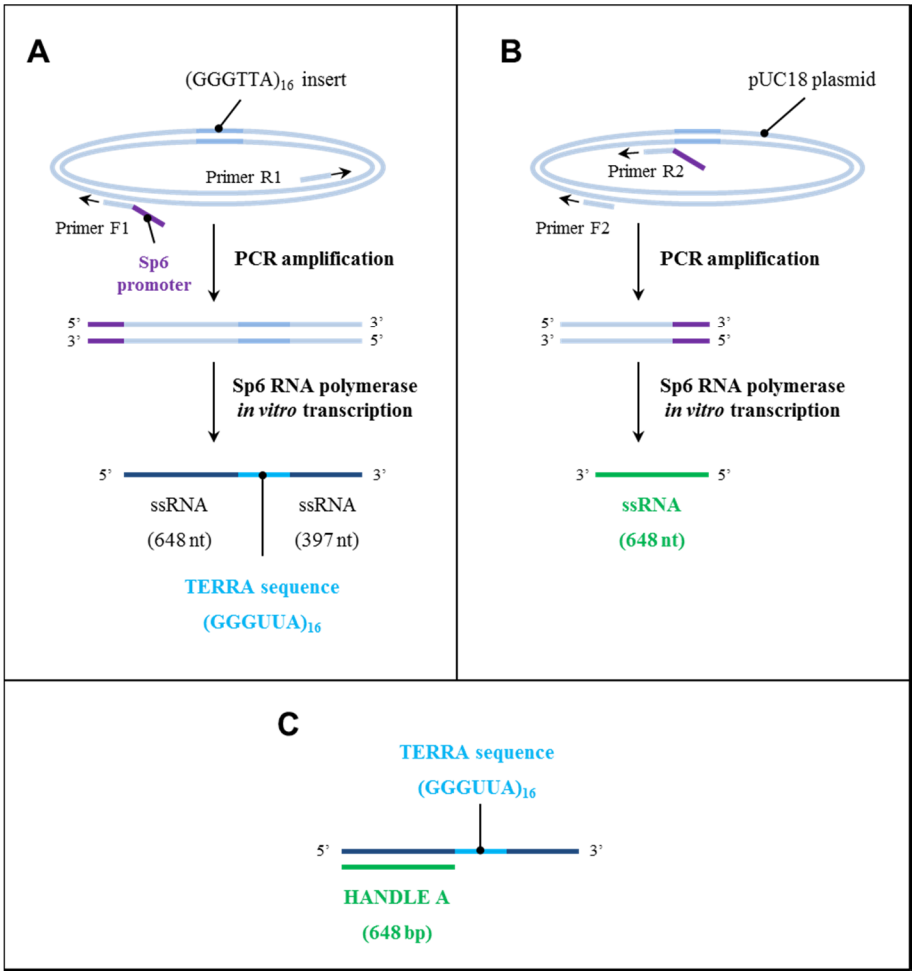

FIGURE 2.3. Schematic representation of the steps involved in the synthesis of TERRA-16 molecules with a single dsRNA handle. **(A)** Synthesis of TERRA sequence; **(B)** synthesis of ssRNA to construct the handle and **(C)** annealing of the two transcripts.

2.2.2. AFM imaging

TERRA imaging was performed using an atomic force microscopy from Nanotec Electronica S.L., operating in tapping mode in air, and using NSG-30 cantilevers (NT-MDT; nominal spring constant, k = 40 N/m; resonance frequency, ν_R = 320 kHz). Under this working mode, the tip is oscillated near its resonance frequency and either the reduction of the oscillation amplitude or the shift in the resonance frequency is kept constant as an image is acquired. This allows imaging molecules weakly attached to a surface in ambient air [255]. Raw topographical images were processed using WSxM freeware [273]. Height measurements were done with respect to the basal salt layer.

2.2.3. Sample preparation for AFM measurements

Before TERRA deposition, the top layer of a mica substrate was cleaved using Scotch tape to reveal an atomically flat and negatively charged surface. Immediately after, it was functionalized with 10 mM $NiCl_2$ during 1 min to allow the subsequent adsorption of the RNA molecules of interest. Excess of $NiCl_2$ was rinsed out with 3 ml of ultrapure Milli-Q-filtered water. Functionalized mica substrate was then dried under a gentle stream of nitrogen gas and checked by AFM to ensure a uniform surface. Finally, TERRA molecules were diluted to a concentration of ~ 0.5 pmol/µl in 10 mM Tris-Cl (pH=7.8) supplemented by 100 mM KCl, deposited onto functionalized mica substrate and incubated at room temperature for 3 min. RNA in excess was removed from the surface using 3 ml of ultrapure water. The surface was then blown dry in a gentle stream of nitrogen gas, being ready for AFM imaging.

2.3. Results

2.3.1. AFM imaging of TERRA-25

First experiments were carried out to observe specific GQ structures in TERRA molecules with twenty-five GGGUUA repeats; i.e. molecules that can give rise to a maximum of six GQs. As shown in *Fig.2.4B*, GQs appear as mountain-shaped protrusions over the mica, with a mean height (± SE) of 1.1 ± 0.03 nm (n= 76, *Fig.2.4D*). Single-stranded RNA overhangs were not visualized, probably because of either RNA degradation or a process of folding around the corresponding GQ structure, forming something like a ball of yarn.

2.3.2. AFM imaging of TERRA-25 with double-stranded RNA handles

To ensure that the previously visualized structures actually correspond to GQs and not to salt agglomerations, molecules were modified in order to have asymmetric dsRNA handles at both ends of the GQ-forming sequence (see *section 2.2.1.2*). As shown in *Figure 2.5 (A)* and (*B)*, topographical AFM images confirm this assembly; GQs appear again as mountain-shaped protrusions over the mica, but this time flanked by fibres of different lengths. Means in height (± SE) of GQ-type structures and dsRNA handles are 0.9 ± 0.03 nm (n=23, *Fig.2.5 (C)*) and 0.5 ± 0.02 nm (n=23, *Fig.2.5 (D)*), respectively. With regards to the lengths of the handles, they were measured using WSxM processor, which allows calculating the total length of a fibre by drawing a series of very short lines along its contour. Then, the mean lengths (± SE) obtained for handle A (648 bp) and handle B (397 bp) were 191.9 ± 1.6 nm (n=23) and 111.4 ± 1.5 nm (n=23), respectively. This results in a nanometre-to-base pair conversion factor of 0.296 ± 0.002

nm/bp for handle A, and of 0.281 ± 0.004 nm/bp for handle B, values consistent with A-family conformations adopted by dsRNA molecules [30, 274].

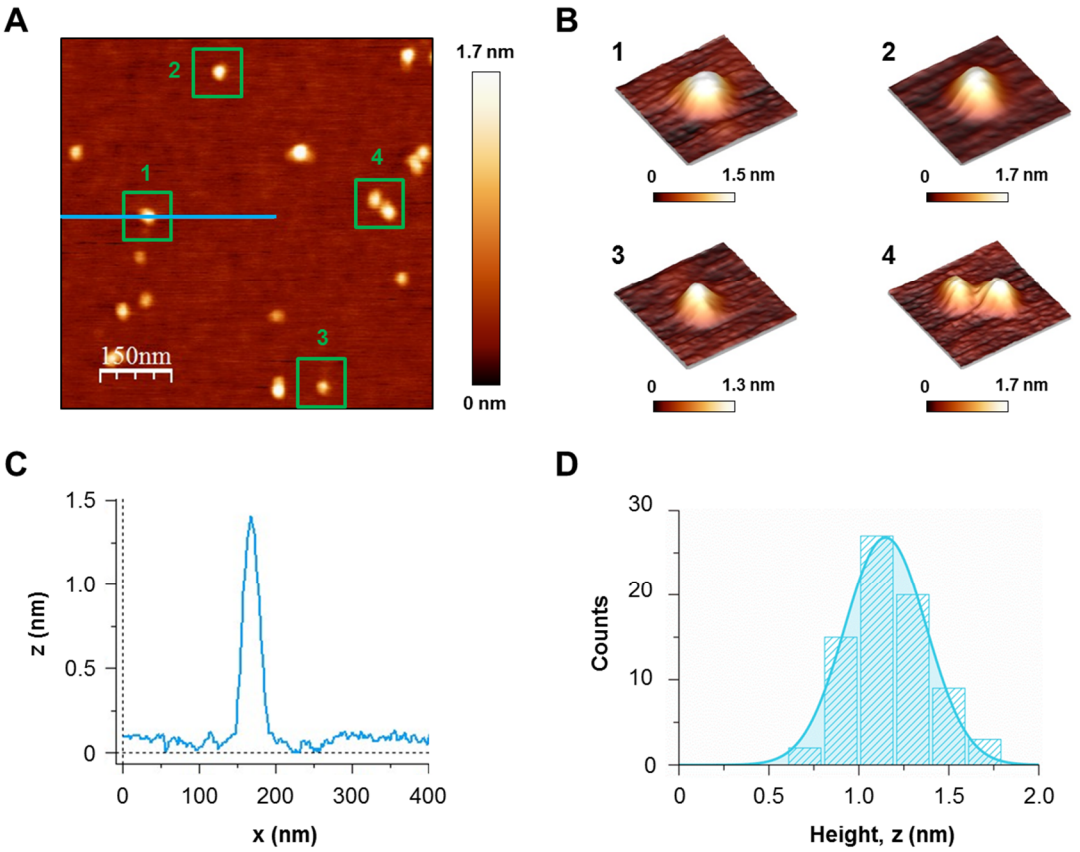

FIGURE 2.4. AFM imaging of TERRA-25 molecules. (A) Dynamic-mode topographical AFM image of TERRA-25 molecules on Ni^{2+} pre-treated mica taken in ambient air. **(B)** Several 3-D topographical AFM images where G-quadruplexes appear as mountain-shaped protrusions over the mica. **(C)** Height profile of a single TERRA molecule. **(D)** Height histogram of the analysed structure (n=76), fitting to a Gaussian distribution (mean ± s.d., 1.1 ± 0.2 nm).

2.3.3. AFM imaging of TERRA-16 molecules

For the purpose of unraveling the geometry that the G-tandems adopt when they are deposited over a surface, TERRA molecules with sixteen telomeric repeats were synthesised; i.e. molecules that can give rise to a maximum of four GQs. In addition, to favour their identification during AFM imaging, the 5' end of these single-stranded molecules was converted into a double helix (see *Fig.2.6 (A)* and *(B)*). Means in height (± SE) of the GQ-type structures and the dsRNA handles are 1.2 ± 0.04 nm (n=28, *Fig.2.6 (C)*) and 0.8 ± 0.03 nm (n=28, *Fig.2.6 (D)*), respectively. The mean length of the handles (± SE) is 188.2 ± 0.9 nm,

given a nanometre-to-base pair conversion factor of 0.0290 ± 0.001 nm/bp, again consistent with A-family conformations adopted by dsRNA molecules [30, 274].

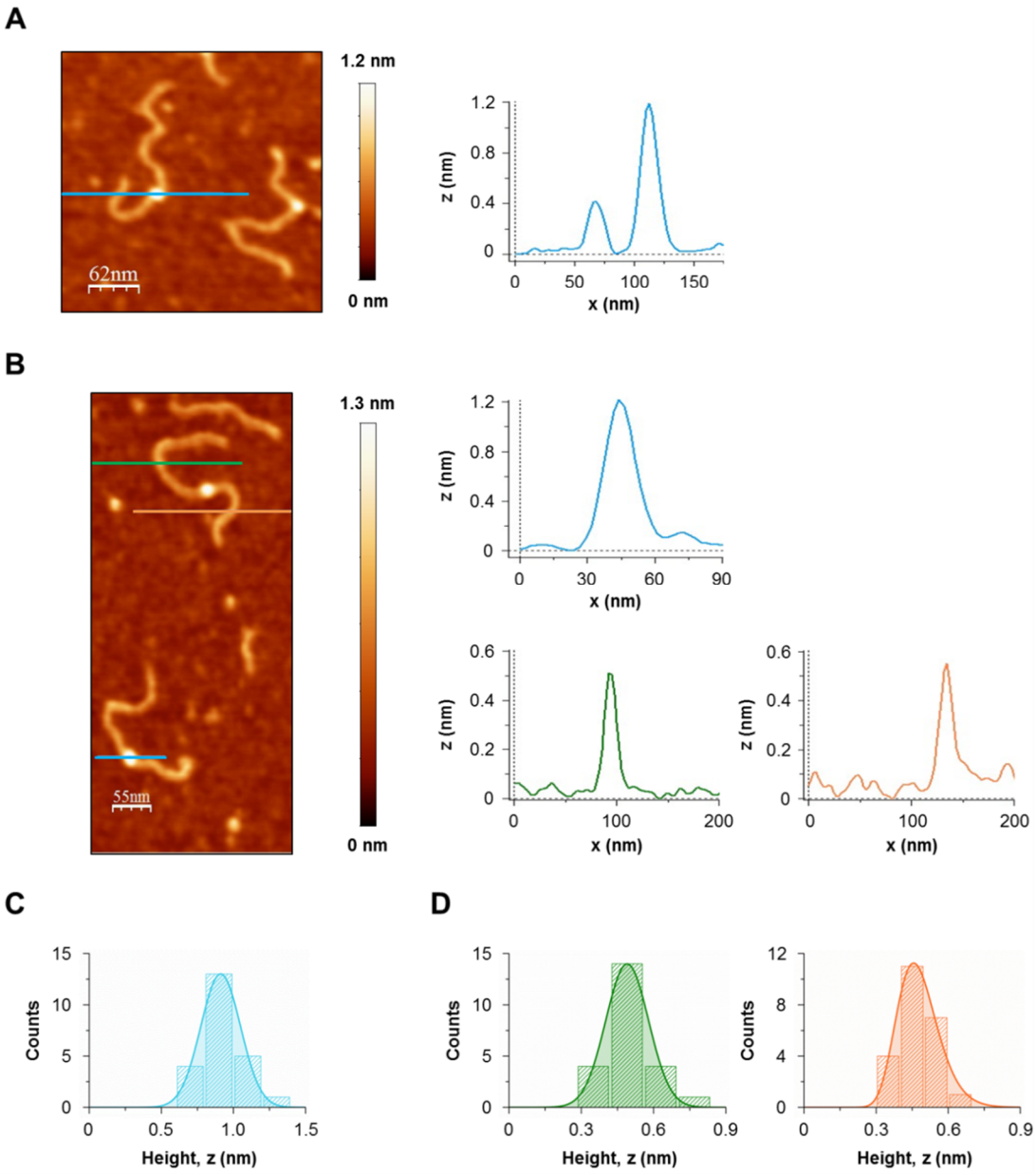

FIGURE 2.5. AFM imaging of TERRA-25 molecules with dsRNA handles. (A, B) Examples of dynamic-mode topographical AFM images taken in air, where height profiles of G-quadruplex structures (light blue) and RNA handles, A (648 bp; green) and B (397 bp; orange), are also shown. **(C)** Height histogram of G-quadruplex structures, fitting to a Gaussian distribution (mean ± s.d., 0.9 ± 0.1 nm). **(D)** Height histograms of dsRNA handles. In green, data corresponding to handle A with Gaussian fitting (mean ± s.d., 0.5 ± 0.1 nm) and, in orange, to handle B with log-normal fitting (median ± s.d., 0.5 ± 0.1 nm).

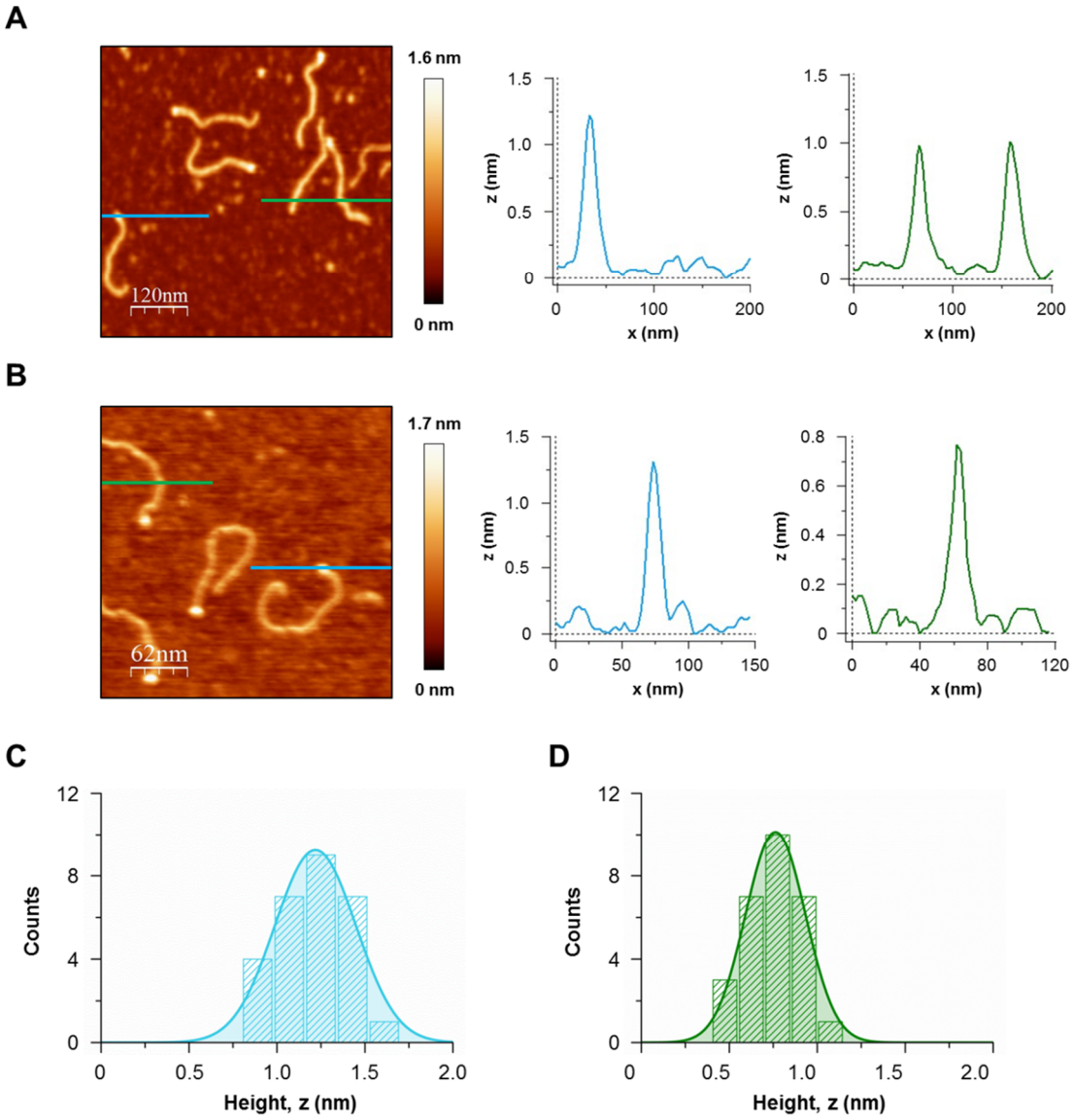

FIGURE 2.6. AFM imaging of TERRA-16 molecules. (A, B) Examples of dynamic-mode topographical AFM images taken in air, where height profiles of G-quadruplex structures (light blue) and dsRNA handles (green) are also shown. **(C)** Height histogram of G-quadruplex structures, fitting to a Gaussian distribution (mean ± s.d., 1.2 ± 0.2 nm). **(D)** Height histogram of dsRNA handles, fitting to a Gaussian distribution (mean ± s.d., 0.8 ± 0.2 nm).

2.4. Discussion

A clear feature in all the obtained AFM images is the appearance of mountain-shaped protrusions. Their identification with tandems of GQs arises from the analysis of the same molecules modified with handles (see *sections 2.3.1* and *2.3.2*), obtaining similar heights on average (1.1 ± 0.03 nm vs. 0.9 ± 0.03 nm). In contrast, by reducing the number of telomeric repeats, no appreciable differences in height were found (1.2 ± 0.04 nm; see *section 2.3.3*), suggesting that consecutive GQs blocks (usually called *G-wires*) could be adsorbed with G-tetrad planes orthogonal to the surface (see *Fig. 2.7)*.

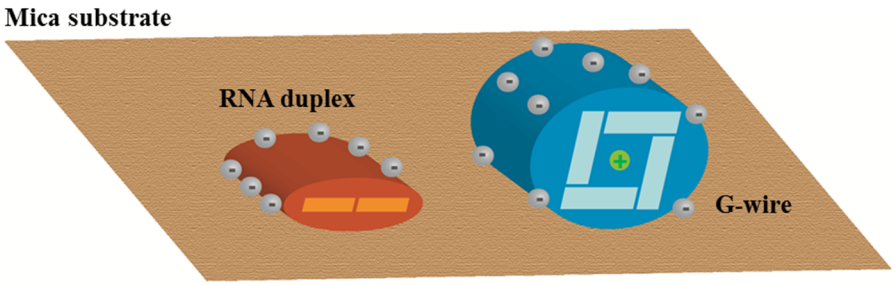

FIGURE 2.7. Elastic deformation of biomolecules during AFM imaging induced by tip–sample interaction. The height of G-wires is not as significantly diminished as in duplex structures, due to the stabilizing capacity of the coordinated cations.

The reduced heights of both GQs structures and dsRNA handles compared with their theoretical dimensions [274, 275] are probably caused by tip-sample interaction during scanning, provoking a strain that compresses the biomolecules [276]. However, the height of handles is significantly diminished compared to the height of the G-wires. A possible explanation could be in the coordinated cations of the quadruple helical structures. They provide an additional stability and, consequently, greater resistance to conformational changes upon surface binding [148].

Chapter 3

The unfolding dynamics of short TERRA molecules

3.1. Introduction

The unfolding of a biomolecule has been majorly studied by chemical denaturation or temperature ramps. However, these methods, which are applied to an ensemble of similar molecules, do not represent the dynamical scenarios in which localized forces are applied by specialised proteins, like the telomerase, to a single molecular structure in the cell. In addition, bulk denaturation methods usually mask the heterogeneity of statistically less populated intermediates by the ensemble averaging process, progressing through pathways with different thermally induced rates relative to those of the mechanical unfolding [277, 278]. Then, single-molecule manipulation techniques have emerged as powerful tools to investigate the movements, forces and strains generated by individual biological species in real time as well as to measure directly multistate or multispecies distributions [168, 194, 196].

G-quadruplex (GQ) unfolding is a halfway problem in complexity between protein and duplex nucleic acid unfolding. It has been proposed that the configurational pathways between the folded and unfolded state comprise a few major intermediate conformations, such as G-hairpins and G-triplexes [279-281]. This is in contrast with the inherent complexity of the protein folding intermediates and the simplicity of a hairpin unzipping. Therefore, G-quadruplex unfolding also represents an intermediate molecular model system from the point of view of non-equilibrium thermodynamics [282].

Single-molecule mechanical folding and unfolding processes of long TElomeric Repeat-containing RNA (TERRA) molecules —susceptible to fold into multiple GQ blocks; see *section 1.2* for a deeper insight— have been studied previously, finding weak cooperativity in the tandem GQ unfolding [221]. In subsequent studies on the human telomeric DNA overhang, a

similar behaviour was found [283]. Yangyuoru et al. analysed, by using a similar single-molecule procedure, a TERRA sequence with only four GGGUUA repeats, which can only fold into a single GQ [284]. These and other authors found that RNA GQs are more stable than previously analysed DNA GQs [224, 285-288].

The next step towards the global understanding of GQ stability, its dynamic processing by motor proteins and its interaction with drugs, is unveiling the influence of neighbouring tracts in the folding pathway of a single TERRA GQ. This chapter will be focused on the study of the unfolding dynamics of short TERRA molecules by using force spectroscopic optical tweezers. In contrast to earlier bulk studies that showed dimorphic competition between a hairpin and a G-quadruplex in either a purposely-prepared RNA sequence [289] or a DNA tract [290, 291], here a single-strand wide conformational competition within a TERRA tract is demonstrated at the single-molecule level.

3.2. Materials and Methods

3.2.1. Synthesis of TERRA molecules

Four different RNA substrates with telomeric GQ-forming sequences were synthesised for their single-molecule manipulation by optical tweezers. In particular, three of them consist of a central ssRNA fragment of *n* repeats (n= 5, 6 or 7) of the telomeric sequence GGGUUA flanked by non-G-rich, random sequences at both sides, UAUA at the 5' end and UAGU at the 3' end (see *Fig.3.1 (A)*). The last one is made up of a central fragment of five GGGUUA repeats flanked by four extra ribonucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3' end (see *Fig. 3.1 (B)*). In all cases, the telomeric RNA sequence together with the flanking extra ribonucleotides are located between two hybrid DNA:RNA handles, which are conveniently labelled for their later attachment to beads through antigen-antibody linkages in the optical tweezers experiments.

By choosing the number of telomeric repeats between five and seven, the formation of TERRA GQ tandems is avoided, since four GGGUUA repeats are needed to give rise to a single conformation of this type. Thereby, the study of the possible influence of neighbouring tracts in the folding pathway of a single TERRA GQ is allowed. The five-repeat RNA molecule with four ribonucleotides at both sides will be considered as the reference system —*TERRA reference construct*. The six and seven-repeat RNA species —*RNA control systems 1* and *2*, respectively— have been designed to analyse the effect of additional telomeric repetitions in the

formation of this helical structure, whereas the five-repeat RNA molecules with a longer random sequence overhang —*RNA control system 3*— will be used to study the influence of adjacent heterogenous ssRNA over a telomeric GQ-forming sequence. The synthesis protocol of each of them is detailed below.

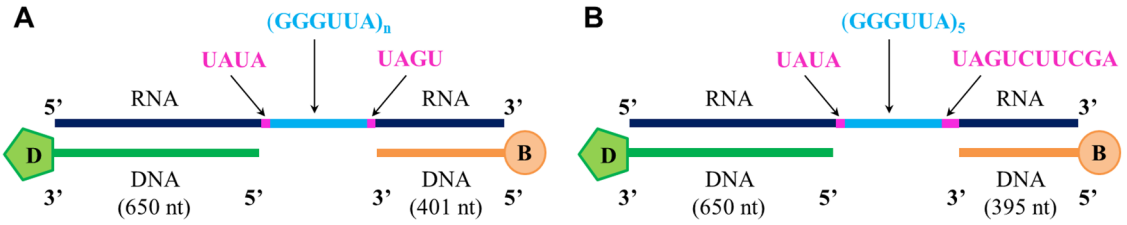

FIGURE 3.1. Scheme of RNA molecular constructions. (A) RNA substrates with a central single-stranded fragment of *n* repeats (n = 5, 6 or 7) of the telomeric GQ-forming sequence GGGUUA flanked by non-G-rich, random sequences at both sides, UAUA at the 5' end and UAGU at the 3' end. In all cases, the complete single-stranded sequence is linked to two hybrid DNA:RNA handles, one of 650 bp at the 5' end (handle A) and the other of 401 bp at the 3' end (handle B), conveniently labelled with digoxigenin (represented as D) and biotin (represented as B), respectively, for their later attachment to beads in the optical tweezers experiments. **(B)** RNA substrate with a central single-stranded fragment of five GGGUUA repeats flanked by four extra ribonucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3' end. The complete single-stranded sequence is also linked to two hybrid DNA:RNA handles, but this time handle B is shorter (395 bp).

3.2.1.1. Synthesis of TERRA reference construct and RNA control systems 1 and 2

The synthesis protocol of these three RNA constructions involves the following steps:

1. Production of modified pUC18 plasmids.

pUC18 plasmids with inserts of five, six and seven GGGTTA repeats were obtained by extraction from E.coli DH5α transformant cells [221] using the *Wizard® Plus SV Minipreps DNA Purification System* (Promega, Ref. A1330), being subsequently sequenced by STAB VIDA to crosscheck the presence of these telomeric repetitions.

2. PCR amplifications of telomeric GQ-forming sequences.

The DNA templates for these TERRA substrates (n=5, 1112 bp; n=6, 1118 bp; n=7, 1124 bp) were obtained by PCR amplification of conveniently modified pUC18 plasmids —with an insert of five, six or seven GGGTTA repetitions as required— using the following pair of primers:

F1: 5'-Sp6 promoter-GAATAAGGGCGACACGGAAATGTTG-3'
R1: 5'-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the number 1 refers to TERRA. The reaction mixture (50 µl) and PCR cycling conditions used here are detailed in *Table 3.1* and *Table 3.2*, respectively. The obtained reaction products were purified using the *Wizard® SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting each sample in 50 µl of nuclease-free water.

3. PCR amplifications of the DNA templates for the hybrid handles.

The DNA template for the hybrid handles, handle A (650 bp) and handle B (401 bp), were also obtained by PCR amplifications of the conveniently modified pUC18 plasmids using the following two pairs of primers:

F2: 5'-GAATAAGGGCGACACGGAAATGTTG-3'
R2: 5'-GTGAGTCGTATTAAGCTTGGCACTGG-3'

F3: 5'- CTTCGAATTCGTAATCATGGTCATAGCTGT-3'
R3: 5'- Biotine-TTTACGGTTCCTGGCCTTTTGCTG-3'

where F stands for forward primer and R for the reverse one, and the numbers 2 and 3 refers to handle A and handle B, respectively. The reaction mixture (50 µl) and PCR cycling conditions are the same as those used for the synthesis of the templates of TERRA sequences (see *Table 3.1* and *Table 3.2*). The obtained reaction products were purified using the *Wizard® SV Gel and PCR Clean-Up System* (Promega, Ref. A9281), eluting each sample in 50 µl of nuclease-free water.

Reagent	**V(µl)**
Ultrapure Milli-Q-filtered water	22.0
Modified pUC18 plasmid (10 ng/µl)	1.0
Forward Primer (Integrated DNA Technologies, 10 µM)	1.0
Reverse Primer (Integrated DNA Technologies, 10 µM)	1.0
SapphireAmp® Fast PCR Master Mix (Takara ClonTech, Ref. RR350A)	25.0

TABLE 3.1. PCR mixture for DNA telomeric GQ-forming sequence amplification.

Cycle		Temperature (ºC)	Time (min)
Hot Start Automatic		94.0	3
Start cycle (x 40)	Denaturation	95.0	0.5
	Annealing	55.0	0.5
	Elongation	72.0	1.5
Temperature		72.0	5
Store		4.0	∞

TABLE 3.2. PCR cycles for DNA telomeric GQ-forming sequence amplification.

4. Labelling of the handles.

Handle A was labelled at its 3' end with a very short tail of digoxigenin-dUTPs using the *DIG Oligonucleotide Tailing Kit, 2nd Generation* (Roche, Ref. 3353583910) at 37ºC during 15 min. Then, the sample was purified using the *Wizard SV Gel and PCR Clean-Up System* (Promega, Ref. A9281). Handle B was labelled with a biotin using a 5'-biotinylated primer (primer R3) during its PCR amplification.

5. Synthesis of ssRNA sequences.

The ssRNA sequences were synthesised by in vitro transcription of the double-stranded (ds) DNA obtained previously using Sp6 RNA polymerase (New England Biolabs; 3h, 40ºC), whose promoter sequence was incorporated at the 5' end of the forward PCR primer (primer F1). The obtained transcripts were purified using RNeasy MinElute Cleanup Kit (Qiagen, Ref. 74204), eluting the samples in 20 µl of RNase-free water.

Reagent		V(µl)
RNase-free water		14.80
Sp6 polymerase buffer 10X (NEB, Ref. M0207S)		2.50
rNTPs dil 1:10	A	1.25
	C	
	G	
	U	
DNA template (~50-60 ng/µl)		1.20
RNase Inhibitor (NEB, Ref. M0314S)		0.50
Sp6 polymerase (NEB, Ref. M0207S)		1.0

TABLE 3.3. Mixture reaction for in vitro transcription of DNA telomeric GQ-forming sequences.

6. Annealing.

The final assembly is achieved mixing equimolar amounts of the corresponding ssRNA and the two DNA labelled handles and annealing them in the presence of 100 mM potassium

chloride, which favours the GQ stabilization, in a PCR apparatus following the procedure specified in *Table 3.4*. A schematic representation of the whole process is shown in *Figure 3.2*.

Cycle	Temperature (ºC)	Time (min)
Denaturation	90.0	5
Annealing	85.0	10
	62.0	90
	52.0	90
	20.0	∞

TABLE 3.4. Annealing procedure of TERRA assemblies for optical tweezers experiments.

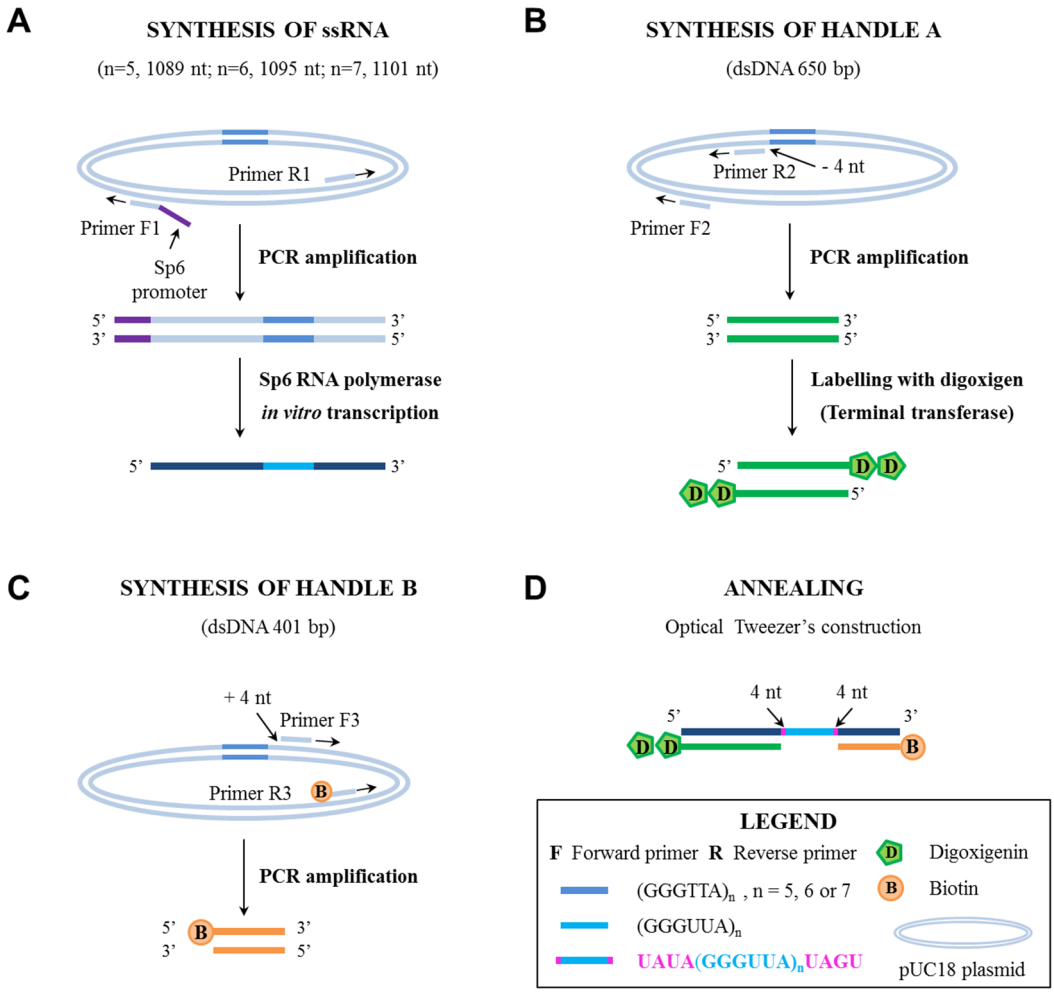

FIGURE 3.2. Schematic representation of the steps involved in the synthesis of the TERRA reference contruct (n=5) and the RNA control systems 1 (n=6) and 2 (n=7) for optical tweezers experiments. The single-stranded (ss) DNA of the 5' hybrid handle was labelled at its 3'end with a very short tail of digoxigenin-dUTP (represented as D) whereas the ssDNA of the 3' hybrid handle was labelled during the PCR amplification using a 5'-biotinylated (represented as B) primer (primer R3). The primers flanking the telomeric repeats, R2 and F3, start at -4 and +4 nucleotides away from the telomeric repeats, respectively. The final molecular construct has as a result extra ssRNA on both sides of the sequence $(GGGUUA)_n$, being n= 5, 6 or 7 (represented in pink in (D)).

3.2.1.2. Synthesis of RNA control system 3

The synthesis protocol of the *RNA control system 3* — i.e. molecules with a central ssRNA fragment of five GGGUUA repeats flanked by four extra nucleotides (UAUA) at the 5' end and by ten extra nucleotides (UAGUCUUCGA) at the 3'end— follows the same steps described in the previous subsection except for the hybrid handle B (see *Fig. 3.3*), which is now shorter (395 bp) to release more random ssRNA at the 3' end of the central fragment (ten ribonucleotides instead of four). Then, the forward primer F3 is substituted by F3', whose sequence is: F3': 5'-ATTCGTAATCATGGTCATAGCTGTTTCCTG-3'

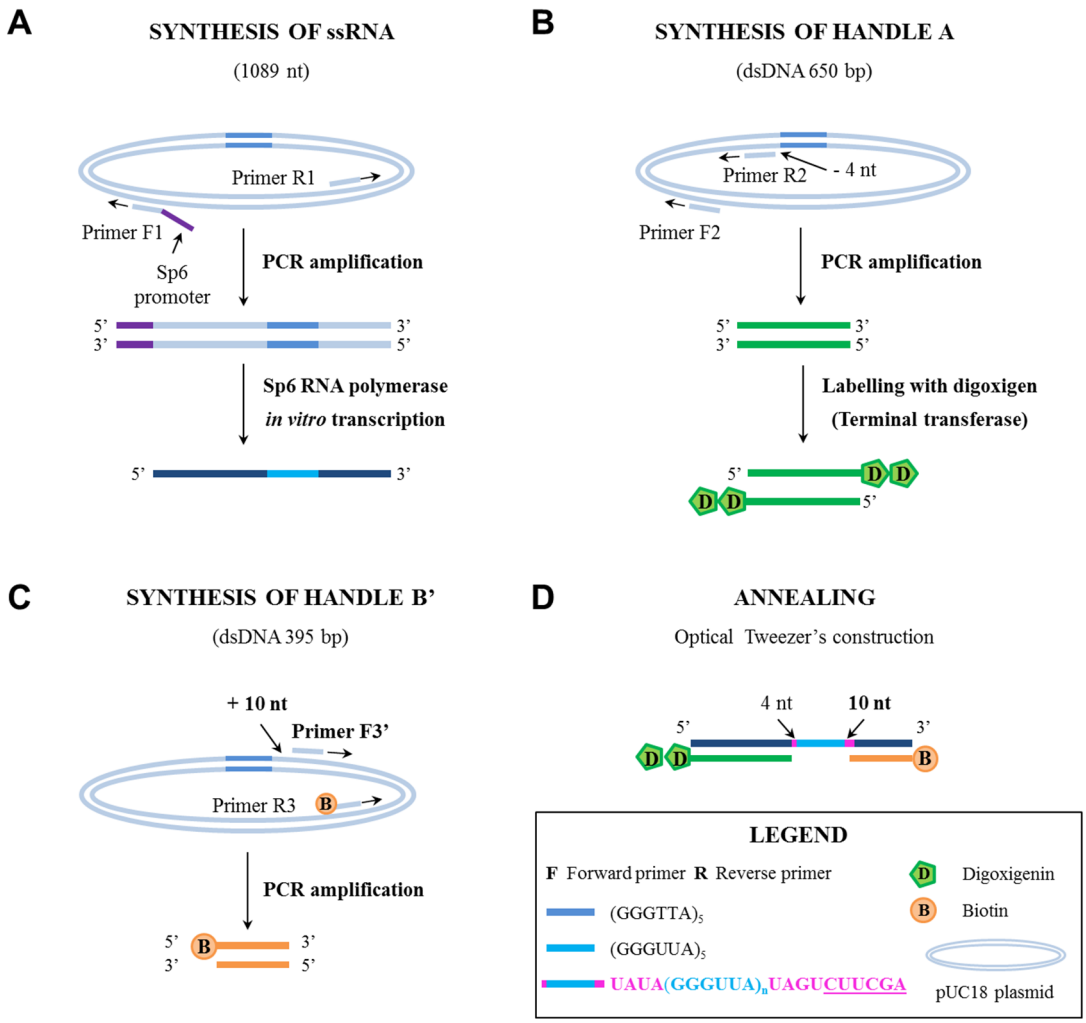

FIGURE 3.3. Schematic representation of the steps involved in the synthesis of the RNA control systems 3 for optical tweezers experiments. The same protocol shown in *Figure 3.2* was followed but substituting forward primer F3 by F3', which releases 10 nt of ssRNA instead of 4 nt.

3.2.2. Synthesis of DNA analogue molecules

DNA analogue molecules to the TERRA reference construct were prepared for comparison purposes. The central ssDNA fragment consists of five repeats of the telomeric sequence GGGTTA flanked by four extra nucleotides at both sides: TATA at the 5' end and TAGT at the 3' end. The handles are made up of two complementary DNA strands with the same sequences as those used in the TERRA hybrid handles. As the transcription step is not needed now, the forward primer F1 does not include the Sp6 promoter at its 5' end. The rest of the protocol follows the same steps as previously described in *subsection 3.2.1.1*. A schematic representation of the whole process is shown in *Figure 3.4*.

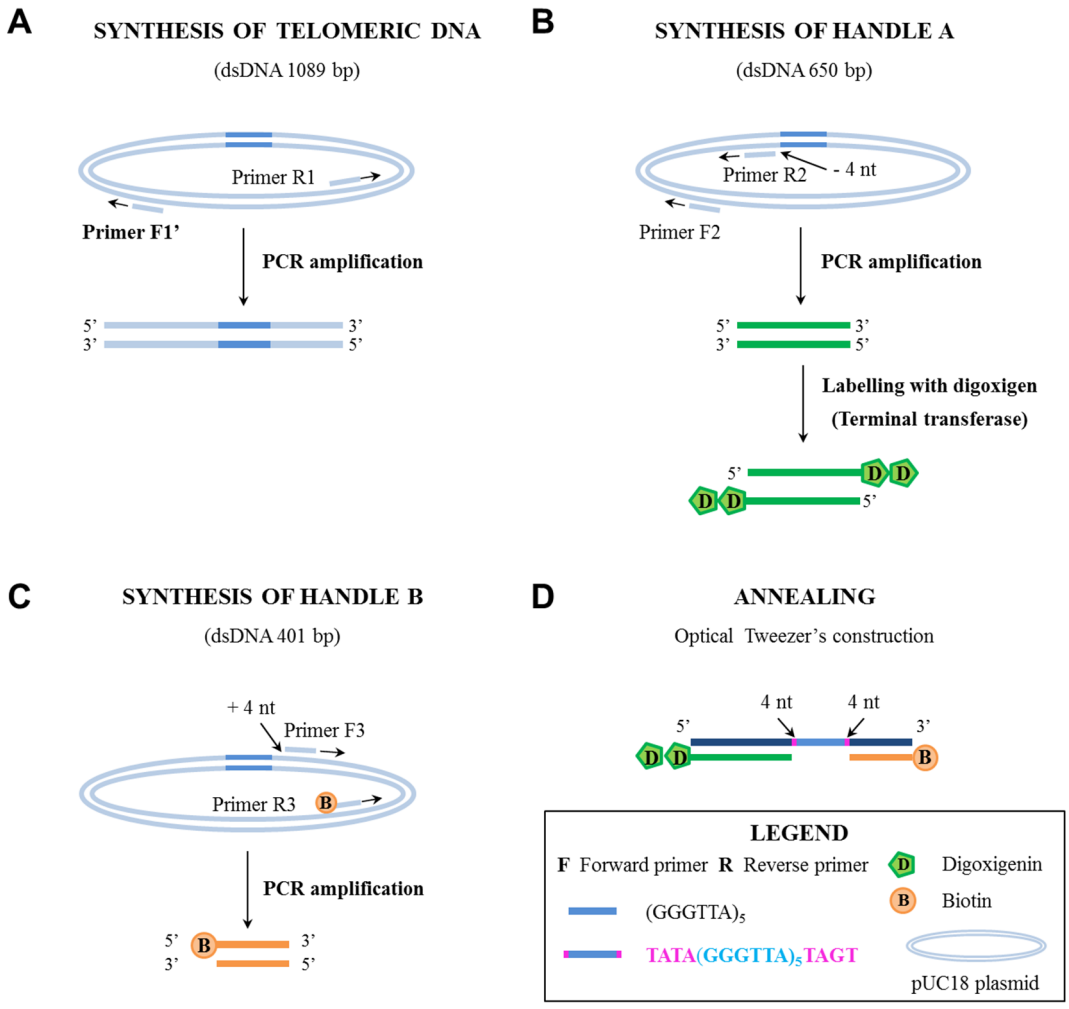

FIGURE 3.4. Schematic representation of the steps involved in the synthesis of the DNA analogous molecules to TERRA reference construct for optical tweezers experiments. The same protocol shown in Figure 3.2 was followed but without the Sp6 promoter at the 5' end of the forward primer F1 (now renamed as F1' in (A)).

3.2.3. Single-molecule experiments

3.2.3.1. Optical Tweezers setups

Measurements have been performed using two optical tweezers (OT) setups with different measurement stability and subsequent resolution, which further served to crosscheck results. The first one is a table-based instrument (*Fig.3.5*), whereas the second one, so-called miniTweezers (mT, *Fig.3.6*), is a compact instrument suspended from the ceiling with reduced mechanical drift and larger measurement stability over time due to the reduction of optical path lengths [292, 293].

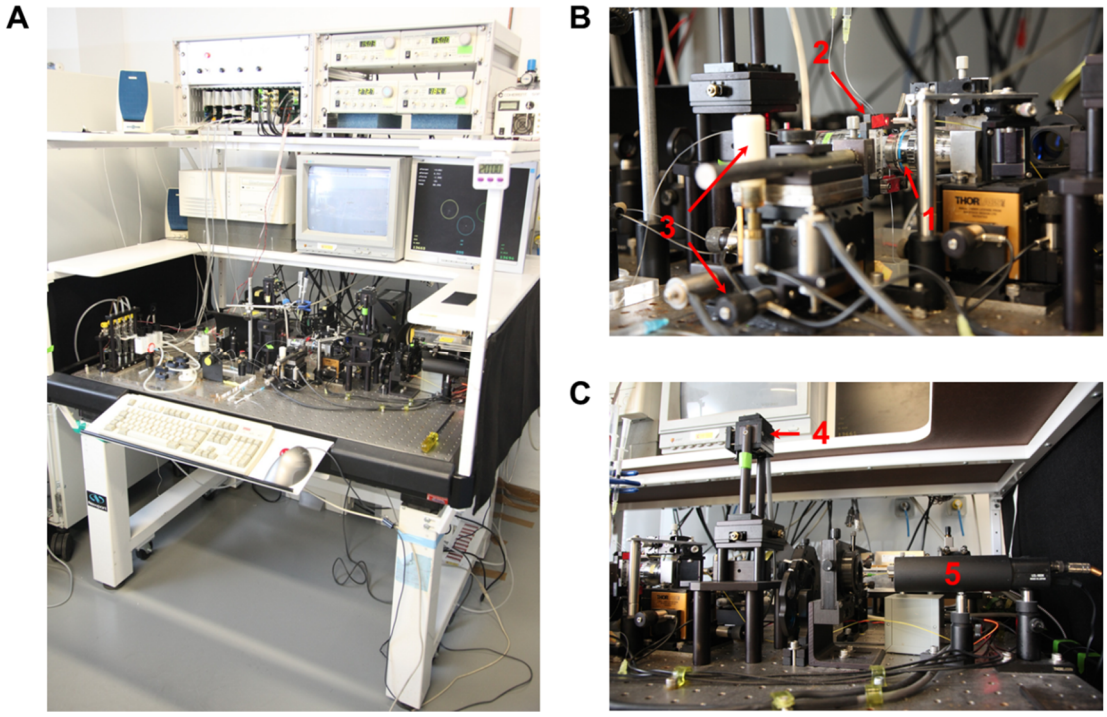

FIGURE 3.5. Table-based OT instrument used in this thesis to perform force-induced unfolding experiments. (A) Frontal view of the whole set-up. **(B)** Detail of the central part of the instrument in which optical trapping is performed, being highlighted one of the water-immersion objectives (1), the microfluidics chamber (2) and the piezo-electric actuators that allow the displacement of the chamber (3). **(C)** Detail of the right side where one of the PSDs (4) and the CCD camera (5) is located.

Beyond these differences, both instruments are based on the same optical design: two diode lasers (the former, 200 mW at maximum power, 835 nm wavelength from *JDS Uniphase* and the latter, 250 mW and 808 nm from *Lumics*) brought to the same focus with orthogonal polarizations, which allow their optical paths to be separated using polarizing beam splitters. Beams are focused according to a counter-propagating configuration through opposing water-immersion objectives, 60X, NA=1.20 (*Nikon CFI Plan-Apochromat* for the table-based setup

and *Olympus, UPLSAPO 60XW* for the mT) to form the optical trap in a microfluidics chamber, which also contains a micropipette. The light exiting the trap from each beam is collected by the opposite objective lens and redirected to position-sensitive detectors, which monitor the three force components acting on the trapped bead. For the calibration based on the conservation of light momentum, it is necessary that all the light exiting the trap is collected; the numerical aperture of the beams is then reduced to approximately 0.5, which allows the objective lenses to act as both light focusing elements and collectors.

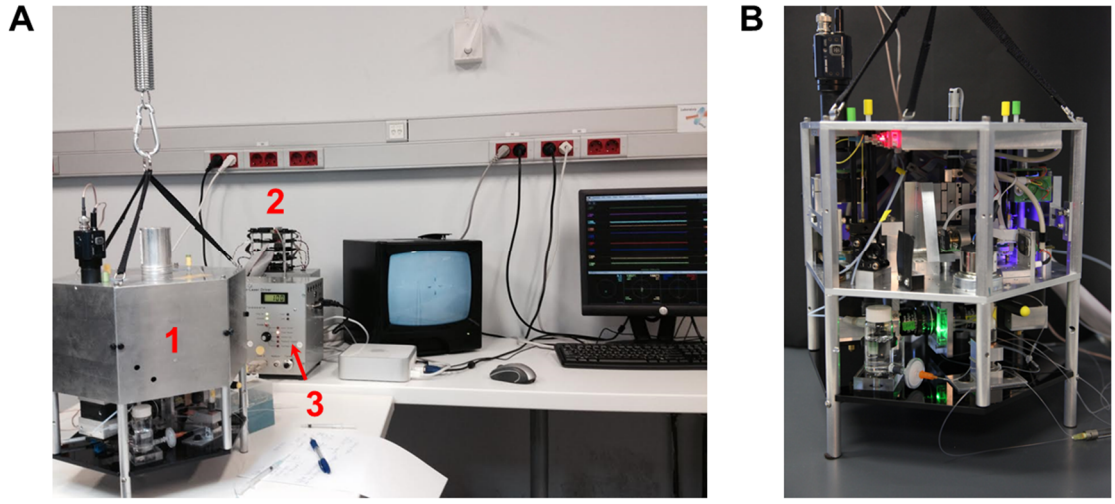

FIGURE 3.6. Minitweezers instrument used in this thesis to perform force-induced unfolding experiments. (A) Frontal view of the whole set-up, where the main parts have been highlighted: (1) head of the mT, which can be suspended from the ceiling by means of a system of pulleys, (2) electronic tower and (3) laser driver box. **(B)** Detail of the interior of the head of the instrument where all the components placed in the optical table in the previous set-up are compacted.

Sample force-extension curves recorded with each instrument are shown along section 3.3 (compare, for example, replicate curves of Fig. 3.9 with those of Figs. 3.14-3.17). Although both instruments gave rise to similar results (approximate force resolution of 0.1 pN and a distance resolution near 0.5 nm), the mT allowed faster valid data collection due to its improved stability. Specifically, force-extension curves, enveloped by thermal noise, were less influenced by artifactual noise enhancing the discernibility of the rupture events associated with the unfolding of the RNA structures.

3.2.3.2. Force-induced unfolding experiments

To perform force-induced unfolding experiments, a single nucleic acid construct made up of a central single-stranded telomeric GQ-forming sequence flanked by duplexes — hybrid in

the case of RNA systems — was tethered by opposite ends between two dielectric polystyrene microspheres: an antidigoxigenin (α-Dig) -coated bead, optically trapped in the laser tweezers, and a streptavidin-coated bead, held by suction on top of a micropipette (see *Fig.3.7*) — located inside of a microfluidics chamber—. Streptavidin-coated, 2.10 μm-diameter beads and G-protein-coated, 2.88 μm-diameter polystyrene beads, purchased from Spherotech (Libertyville, IL), were used. α-Dig was linked to the G-coated beads using dimethyl pimelimidate (DMP) dihydrochloride.

In the table-based instrument, force on the studied molecule was generated by moving the micropipette —fixed to the microfluidics chamber— relative to the stationary optical trap through a piezo-controlled stage. In the miniTweezers, it is created by moving the trap relative to the stationary micropette through bending the lasers' optical fibres using piezoelectric crystals. The extension of the molecule was determined, in both systems, from the distance between the centres of the beads. Stretch-relaxation cycles were performed at 10.2 ± 2.7 nm/s, which corresponds to a pulling rate of 5 ± 1 pN/s near the rupture events. All tweezers experiments were carried out at room temperature (23 ± 2 °C) in 10 mM Tris-HCl (pH 7.8) supplemented by 100 mM KCl and 1 mM EDTA. To avoid RNA degradation, diethylpyrocarbonate (DEPC)-treated water was used (DEPC to 0.1% v/v; incubation at 37ºC overnight followed by autoclaving).

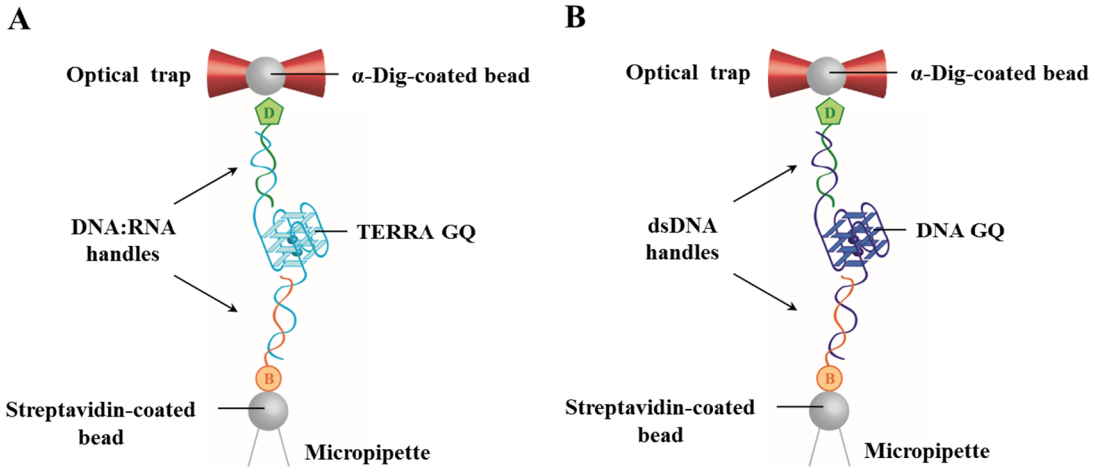

FIGURE 3.7. Experimental set-up for optical tweezers measurements. (A) Experimental configuration for TERRA reference construct and RNA control systems. A single RNA molecular construction was tethered by opposite hybrid duplex ends between two polystyrene microspheres: and α-digoxigenin-coated bead, optically trapped, and a streptavidin-coated bead, held by suction on top of a micropipette. **(B)** Experimental configuration for DNA analogue molecules to TERRA reference construct. In this case, a single DNA molecular construction was attached to polystyrene beads by DNA duplex ends.

3.3. Results

We singled out several repetitions, *n*, of the human telomeric sequence GGGUUA and added extra nucleotides (nts) of non-G-rich, random sequences on both the 3' and 5' ends of ssRNA (see *Materials and Methods* section for further details on the synthesis). We chose $4 < n < 8$, first, to make sure that the sequence can only fold into a single G-quadruplex conformation and, second, to study the effect of additional repetitions in the formation of this tertiary structure.

For reference purposes, we will focus on the five-repeat molecule (n = 5) with four extra nucleotides on both the 3' and 5' ends. Earlier experiments by circular dichroism and NMR with this RNA fragment demonstrated its potential to form an intramolecular parallel G-quadruplex [294]. We will next show that the 8-nt ssRNA excess together with an extra GGGUUA repetition provides ample configurational flexibility to this telomeric sequence to fold into several conformations. In fact, this 38-nt sequence can give rise to: (i) a four-stranded, **G-quadruplex** structure interacting with different lengths of excess ssRNA on both ends (see *Fig.3.8*), (ii) a three-stranded, **G-triplex** similarly coexisting with two ssRNA lateral fragments of variable lengths and (iii) either a **hairpin-like structure or more general condensed conformations of ssRNA**. These structural and subsequent kinetic scenarios as well as the probabilities for the conformations discussed above can only be analysed by real-time, single-molecule experiments because bulk approaches in which an ensemble of equal preparations take place can only give rise to average signals thus hindering the access to the less populated conformations [239, 295].

We have used OT to study single-molecule mechanical folding and unfolding processes under controlled *in vitro* conditions. To this end, the 38-nt ssRNA sequence was linked to hybrid DNA-RNA handles, which opposite ends were attached to polystyrene beads, one optically trapped and the other held by suction on top of a micropipette (see *Fig.3.7(A)*). By moving the micropipette relative to the optical trap, force-extension curves of the molecular construct can be recorded in real time. Rupture events become observable as jumps over the characteristic force-extension curves of the double-stranded hybrid handles, whose elastic response is similar to those of dsDNA and dsRNA [296, 297]. Although not deterministically, conformations can be tracked by force and extension signals and the interference or compatibility among the different conformations can be distinguished by their simultaneous occurrence in single unfolding/refolding traces, as reflected in the force-extension curves and in the further statistical analysis. In this regard, we registered **two types of behaviours** in the

force-extension experiments: **curves with single or two consecutive rupture events**, which will be analysed next.

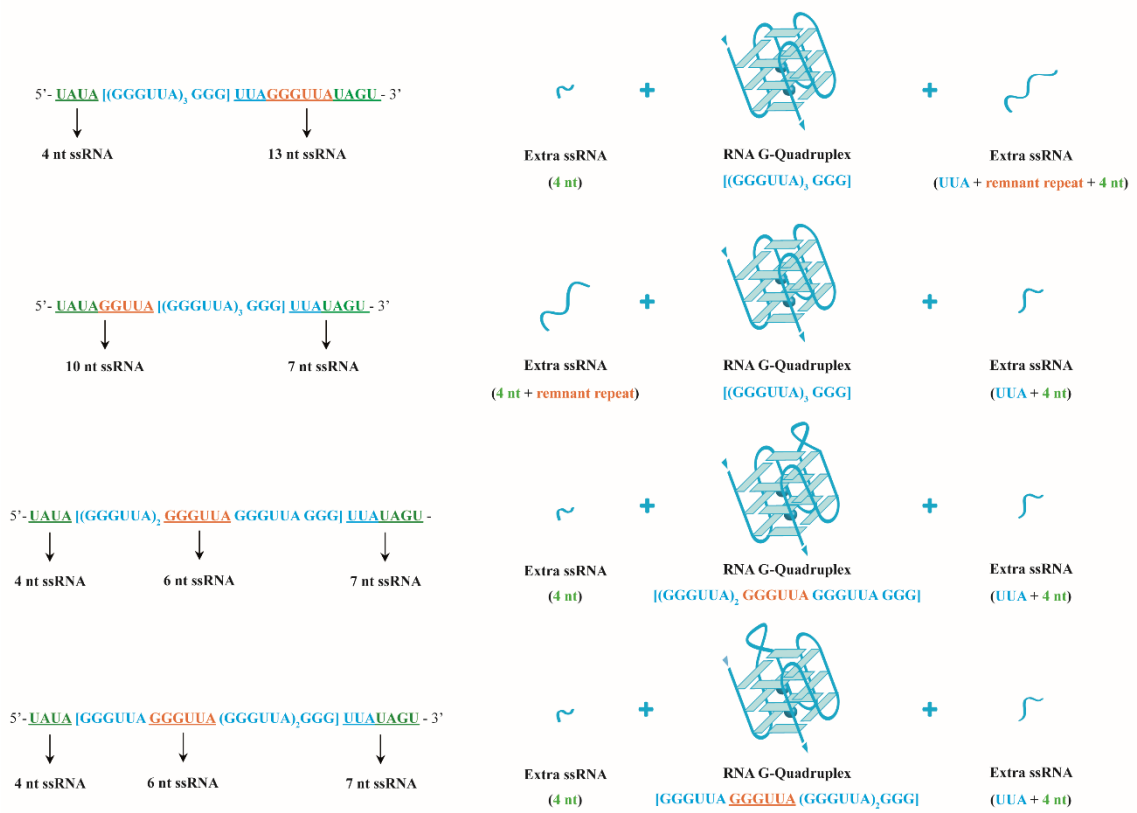

FIGURE 3.8. Possible G-quadruplex assembly configurations for the TERRA reference construct, made up of five GGGUUA repeats with eight random nucleotides spread at the 3' and 5' ends to a total of 38 nucleotides.

3.3.1. Single rupture events

Force-extension curves with only one rupture event represented the typical behaviour. *Figure 3.9* shows the results of several single molecules with the reference sequence. As mentioned above, the traces mostly reveal the characteristic elasticity of the double-stranded handles interrupted by rupture events at intermediate forces. Some of the traces revealed the overstretching transition of the hybrid handles, what confirms that only one molecular construct has been attached between the beads [297]. Rupture signatures correspond to unfolding events of different species, as judged by their force and extension values and also by the relaxation behaviour of the traces (see also insets in *Fig.3.9*).

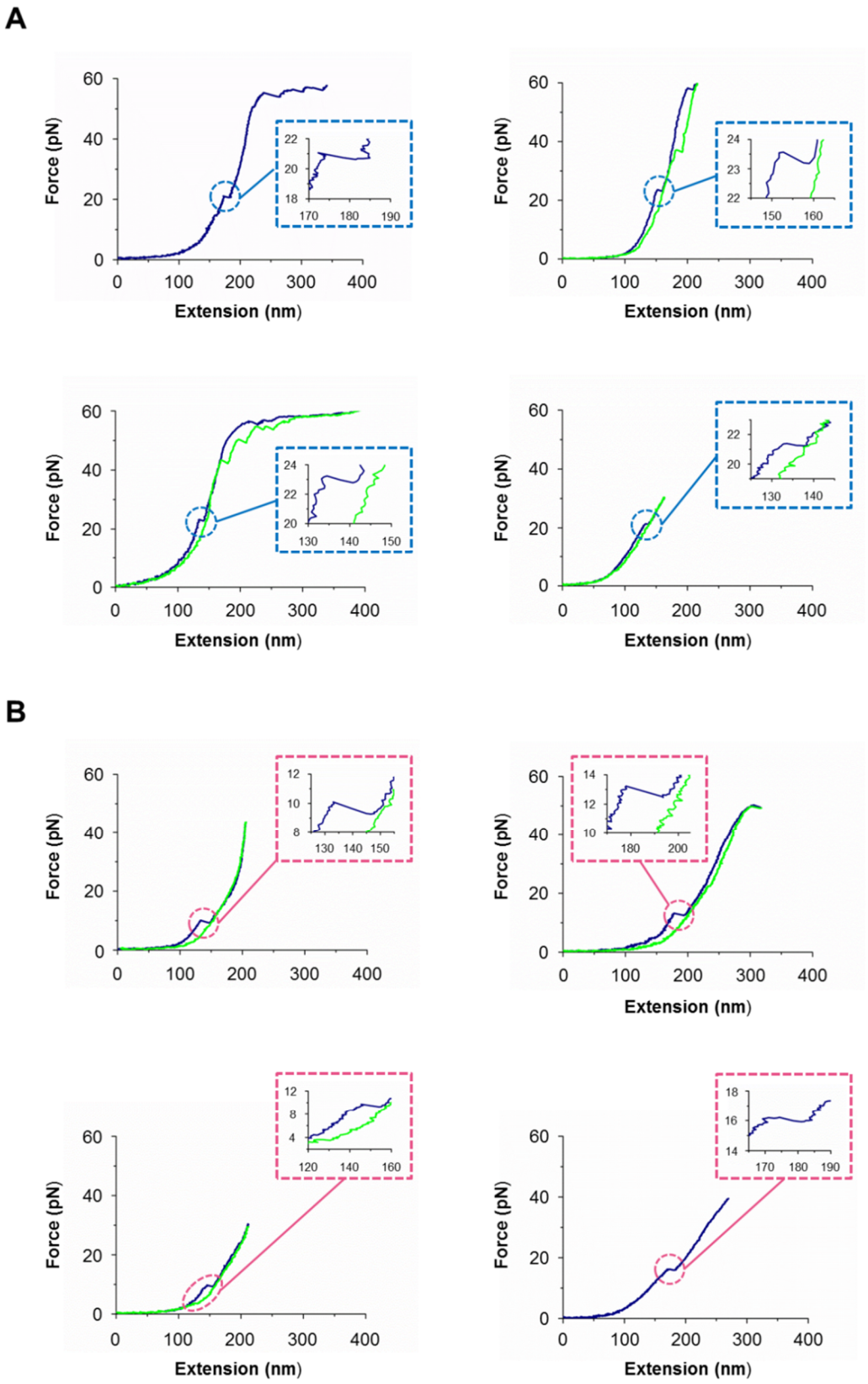

FIGURE 3.9. Replicates of force-extension curves of the TERRA reference construct in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded condensed conformations are probably unfolded and refolded (blue and green lines, respectively).

Three different types of rupture events can be distinguished and they can be assigned, *a priori*, to the unfolding of different structures by estimating the released ssRNA. Let L_{nt} be the mean, effective distance between successive nucleotides in ssRNA, d_G the solution-structure distance between first and last nucleotides in the G-structure (PDB ID: 2KBP) [284, 298], N_G the number of nucleotides in the folded structure and x the rupture extension measured in the OT experiments, then:

$$x = N_G \times L_{nt} - d_G \quad (1)$$

Applying equation (1), the estimated released extension upon the G-quadruplex and the G-triplex ruptures are: 21 (nt) × L_{nt} – 1.5 (nm) = 9.2 nm and 15 (nt) × L_{nt} – 1.9 (nm) = 5.8 nm, respectively, where we have used L_{nt} = 0.51 nm/nt following Yangyuoru et al.[284] (see *Fig. 3.10*). The interphosphate distance actually depends on the force at which the rupture takes place because it is related to the ssRNA elastic response. For ssDNA, Alemany et al.[299] found an effective length across forces of L_{nt} = 0.58 nm/nt according to fittings of the force-extension curves of short ssDNA to inextensible elastic models. Since the inter-phosphate crystallographic distance in DNA (~0.7 nm/nt) is longer than in RNA (~0.59 nm/nt) [296, 297, 300], it is expected that the effective distance between phosphates in the ssRNA under tension complies with the number used above. Then, we can conclude that the most probable rupture events arise from the unfolding of a G-quadruplex/triplex (*Fig.3.9 (A)*) as well as of more general structures caused by ssRNA self-interactions (*Fig.3.9 (B)*). This interpretation agrees with previous studies: G-triplex and G-quadruplex rupture events were located primarily in the range of 20-30 pN [284], whereas single-stranded interactions were noticed between 4-10 pN, depending on ionic conditions and temperature [292, 299, 301-303].

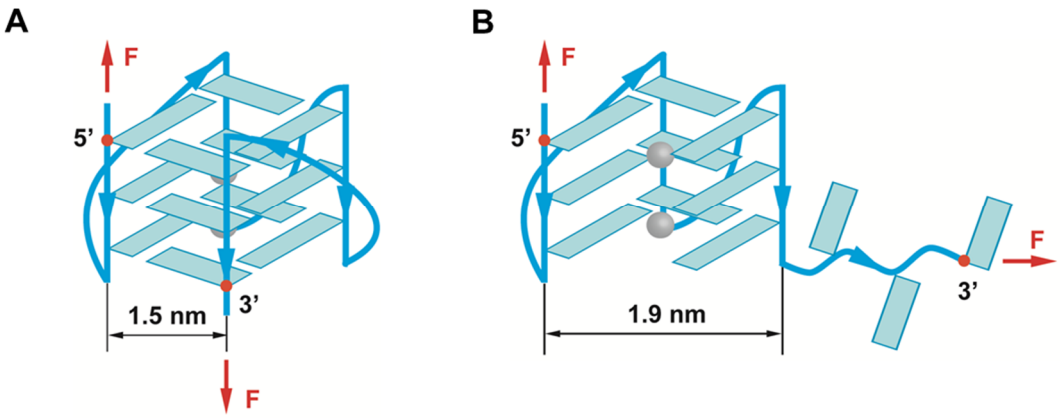

FIGURE 3.10. End-to-end distance measurements of non-canonical conformations. **(A)** TERRA G-quadruplex and **(B)** putative G-triplex-like structure. *F* stands for the tension applied on the structure in each single-molecule experiment. Red arrows represent pulling direction and orange circles the first and last nucleotide of the sequence.

To *in situ* confirm this probabilistic association we recorded data over almost 300 rupture events of this reference molecular construct. In the assays, we sometimes subjected the same molecule to more than one stretch-relaxation cycle, each separated by a pause of 5-10 s in which the force had been relaxed to zero pN. During this pause, the molecule had extra-time to attempt refolding in case it had not happened during the relaxation pathway.

The statistical analysis of the force-extension curves for single-rupture events is presented in *Figure 3.11*. Panel (A) shows the rupture force distribution, where two major populations can be distinguished centred at nearly 9 and 23 pN. According to these forces, the proposed structures as majorly responsible for these two populations would be hairpin-like structure or condensed ssRNA (due to ssRNA self-interactions), for the lower and narrower force peak, and the G-quadruplex/triplex unfolding, for the higher and wider force peak. Panel (B) shows the unfolding extension histogram.

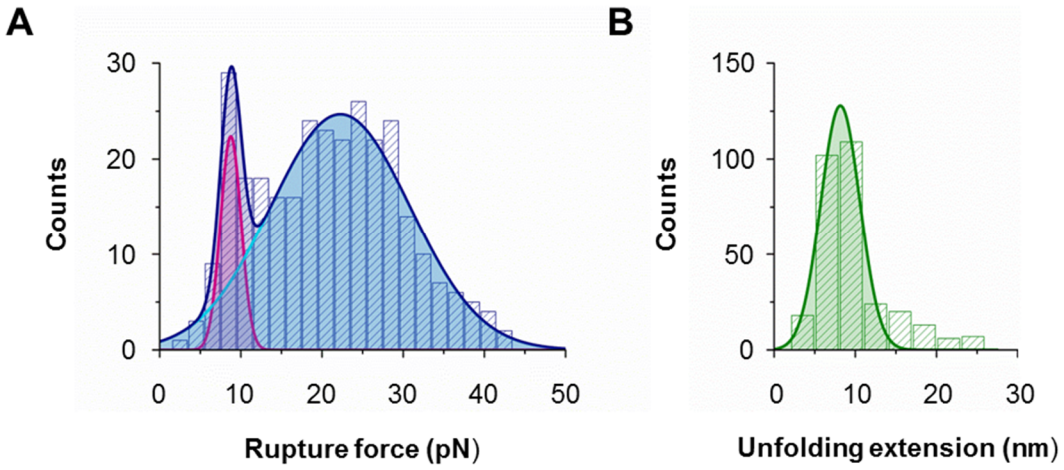

FIGURE 3.11. Statistics of single rupture events in the TERRA reference construct. (A) Rupture force and **(B)** unfolding extension distribution for TERRA molecules with five GGGUUA repeats flanked by four random ribonucleotides at both sides (N=299). Mean ± standard error (s.e.) and standard deviation (s.d.) according to Gaussian fitting are: (A) 8.8 ± 0.2 and 1.3 pN (pink line), 22.3 ± 0.5 and 8.6 pN (light blue line) and (B) 8.2 ± 0.2 and 2.4 nm. In (A), dark blue line corresponds to Gaussian double-peak fitting.

To further confirm the interpretation of these two statistical populations, two filters in extension have been applied, based on the theoretical dimensions of G-triplex and G-quadruplex structures [284]: a low-pass filter with a cut-off extension of 12 nm, which favours rupture events corresponding to the unfolding of the G-quadruplex/triplex conformations, and a high-pass filter above 12 nm, which favours the populations corresponding to the unfolding of hairpins and those with more random behaviour related with ssRNA self-interactions [289-291].

Figure 3.12 shows the resulting force histograms. The populations correlate with the expected conformations. Namely, *Figure 3.12 (A)* shows a broad peak centred at approximately the same force as that shown for the wide population in *Figure 3.11 (A)*, being attributed to the G-quadruplex/triplex unfolding. On the other hand, *Figure 3.12 (B)* populates majorly the low forces, including a peak at 7-8 pN similar to the narrow one in *Figure 3.11 (A)*, which has been associated mainly to the unfolding of generally condensed ssRNA structures.

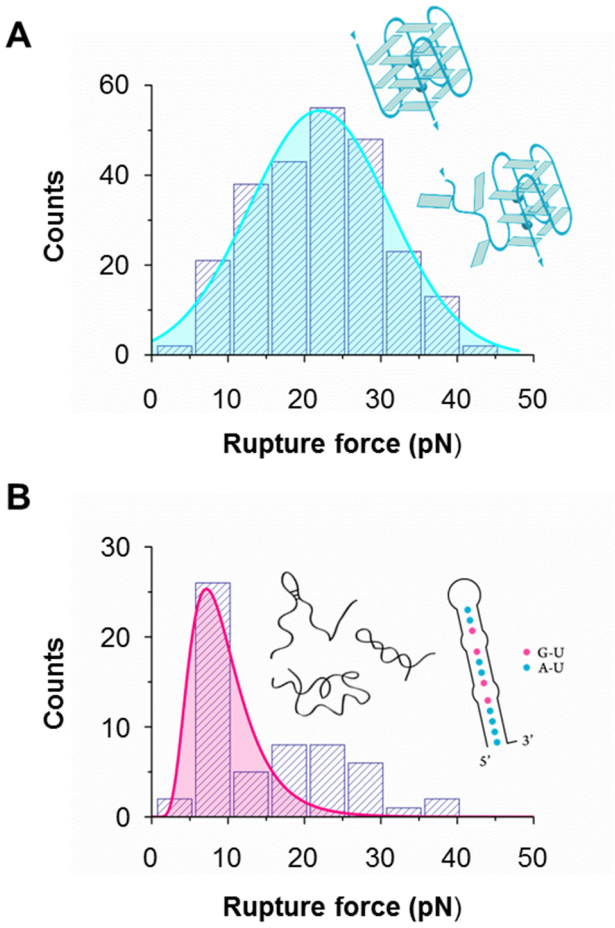

FIGURE 3.12. Rupture force distributions for the different conformations in the TERRA reference construct. (A) G-Quadruplex and G-triplex (N= 241, 80.6%), data filtered in the range [0, 12] nm. Mean ± s.e. and s.d. according to Gaussian fitting is 22.0 ± 0.6 and 9.3 pN. **(B)** Duplex or other condensed conformations (N=58, 19.4%), data filtered in the range (12, 26] nm. Median ± s.d. according to a log-normal distribution is 8.7 ± 4.4 pN. Peak value (mode) ± s.e. is 7.2 ± 1.2 pN.

The G-triplex has been described as an intermediate, possibly in-pathway conformation in the folding/unfolding of the G-quadruplex with similar stability to this four-stranded structure [279-281, 287, 304-307], although both conformations may not be resolved within the same single unfolding pathway [284, 308-310]. These forces for the G-quadruplex/triplex are comparable within experimental error to those reported in [284], as expected, since we used a similar loading rate (see *Materials and Methods*), hence working similarly away from

equilibrium [285, 311]. They are also higher than those found in [294] because in that study rupture events were not classified but considered within the same statistics hence giving rise to a mean force between those of the two populations distinguished in *Figure 3.11 (A)*.

With regards to the peak at low force, apart from the randomly-condensed conformation [292, 299, 301, 302, 312, 313], a very weakly stable hairpin may be formed (see insets in *Fig.3.12 (B)*). This duplex possesses two G-bubbles and G-U and A-U pairings in the stem, according to mfold simulations with the 5-repeat reference sequence (ΔG= –0.40 kcal/mol at 1 M NaCl) [314]. This duplex does not correspond to the G-hairpin, described as a possible G-triplex or G-quadruplex intermediate that results from two repetitions of the GGGUUA sequence [279-281]. It extends over almost the entire ssRNA sequence but its free energy does not supersede that of the thermal fluctuations (–0.7 k_BT, being k_B the Boltzmann constant and T the temperature), which yields an equilibrium rupture force much lower than 1 pN. This makes the duplex very improbable according to the rupture force distribution of *Figure 3.11 (A)*.

3.3.1.1. Comparison with the DNA analogous system

Figure 3.13 shows the statistics of unfolding/refolding experiments with an equivalent molecular construct to the reference system entirely made of DNA, thus preserving sequence and length of both the central single-stranded segment and the double-stranded handles for attachment to beads (see *Fig. 3.4* and *Fig.3.7 (B)*). This molecule has also a potential to fold into a G-quadruplex, as reflected in the force-extension curves represented in *Figure 3.14*.

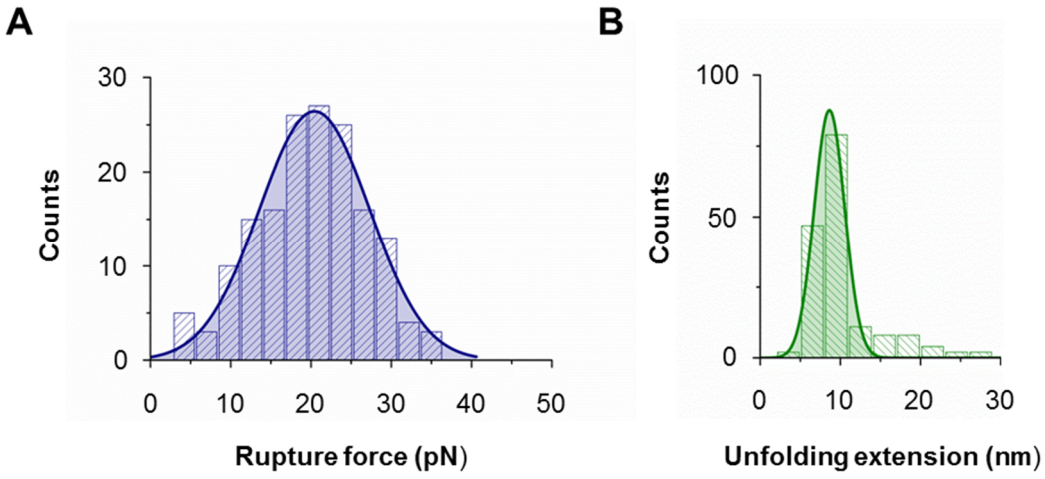

FIGURE 3.13. Statistics of single rupture events in the DNA analogous system to the TERRA reference one. **(A)** Rupture force and **(B)** unfolding extension distribution for DNA molecules with five GGGTTA repeats flanked by four random nucleotides at both sides (N=163). Mean ± standard error (s.e.) and standard deviation (s.d.) according to Gaussian fitting are: (A) 20.4 ± 0.4 and 6.9 pN and (B) 8.6 ± 0.1 and 1.9 nm.

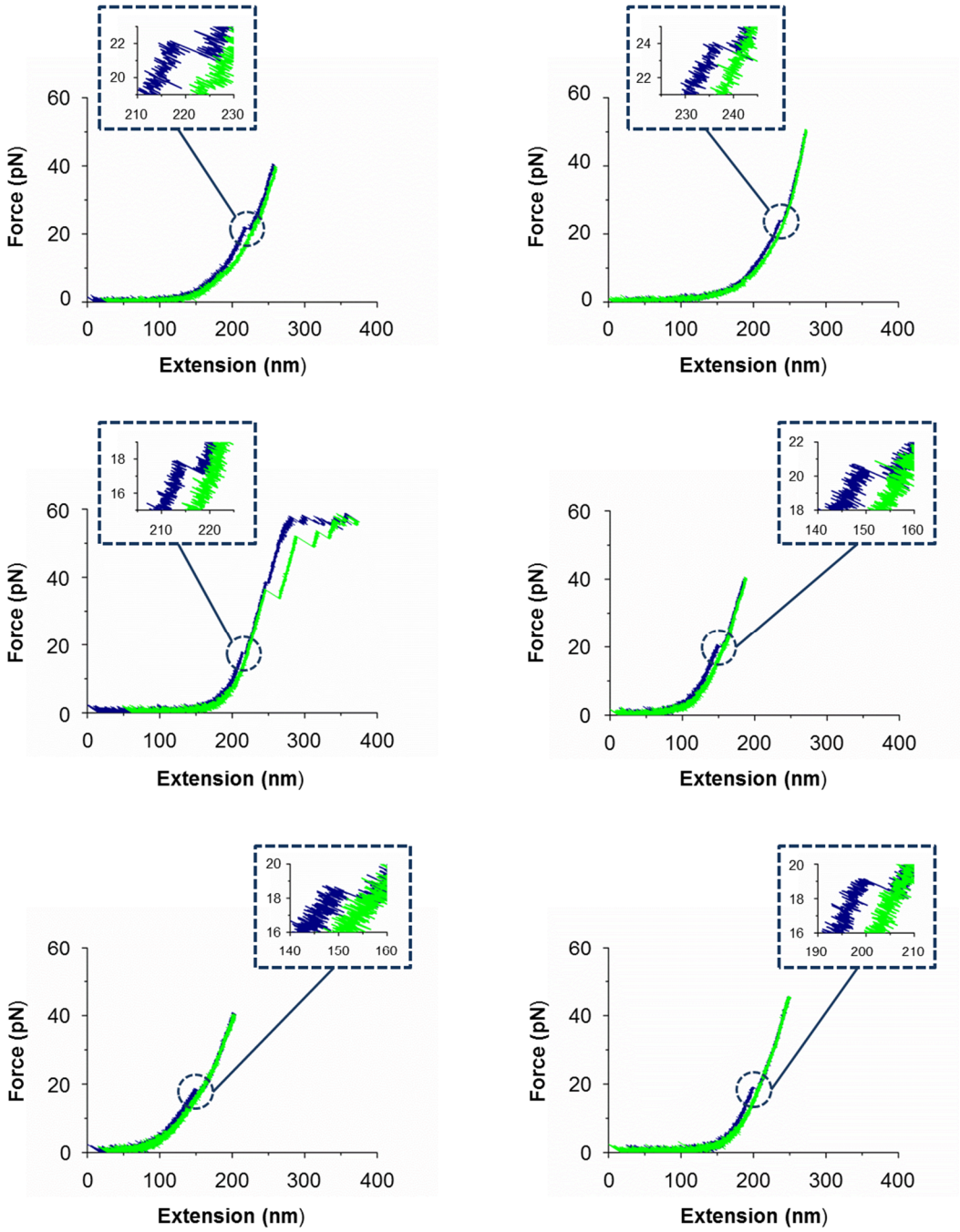

FIGURE 3.14. Replicates of force-extension curves of DNA molecules with five GGGTTA repeats in which a G-quadruplex/triplex is probably unfolded and refolded (blue and green lines, respectively).

Its rupture force statistics (*Fig.3.13 (A)*) exhibits a single broad peak centre at around the same force of that corresponding to the G-quadruplex/triplex. The peak force is shifted to the left by less than 2 pN with respect to the RNA distribution corresponding to the G-quadruplex/triplex (*Fig.3.12 (A)*), in agreement with previous works that reported that DNA G-quadruplex unfold at lower force [224, 285-288].

The most important feature of these control experiments is the absence of a peak near 9 pN, which were crosschecked with measurements in a high-resolution instrument (see *Materials and Methods* section and compare curves from *Fig.3.9* and *Fig. 3.14*). It indicates that the potential of the RNA molecule to self-fold and thus block highly-organized structures is unique. In fact, stronger hybridization interactions in RNA than in DNA of equal sequence have been observed both in bulk and in single-molecule experiments previously [297]. Beyond electrostatic base-pairing considerations, RNA possesses higher tendency to self-interact due to the puckering and chemistry of the bases [300, 315], thus making G-quadruplex sequences in RNA more prone to exhibit conformational competition in the presence of extra single-stranded random sequences than in DNA. Moreover, the weakly stable hairpin that can be adopted by the TERRA reference tract is not possible in the analogous ssDNA ($\Delta G > 0$), reflecting again the higher conformational potential of RNA relative to ssDNA [300].

3.3.1.2. Comparison with the RNA control systems

To study the effect of both additional repetitions and the sequence of the neighbouring tracts in the formation of a single TERRA GQ, three other RNA molecular constructs were synthesised: a 6-repeat and a 7-repeat molecules flanked by the same random sequences as previously used (44 and 50 nts, respectively) and a 5-repeat construction with more extra random RNA sequence (44 nts) (see *Materials and Methods* section). The qualitative behaviour is maintained for single rupture events, as reflected in their force-extension curves (*Fig. 3.15-3.17*): rupture events at both high forces, compatible with G-quadruplex/triplex, and low forces are observed.

The statistics over convergent populations of experiments are shown in *Figure 3.18*. For the 6-repeat molecule, *Figure 3.18 (A1)*, two populations are distinguished, being the high-force distribution centred at a similar force (within experimental error) than that of the 5-repeat molecule previously studied (*Fig.3.11 (A)*). However, the low force distribution presents this time a shift to higher forces, approaching the high-force peak compatible with the G-quadruplex/triplex conformation. In agreement with this trend, this low-force distribution overlaps with that of the high-force for the 7-repeat molecule (see *Fig.3.18 (B1)*). For the case of the 5-repeat molecule with extra random sequence, *Figure 3.18 (C1)*, it is observed again the bimodal distribution shown in *Figures 3.11 (A)* and *3.18 (A1)*, with forces for the low and high force peak similar to those of the 6-repeat molecule (*Fig. 3.18 (A1)*). This is expected since the number of nucleotides involved in these two molecules is identical (44 nt).

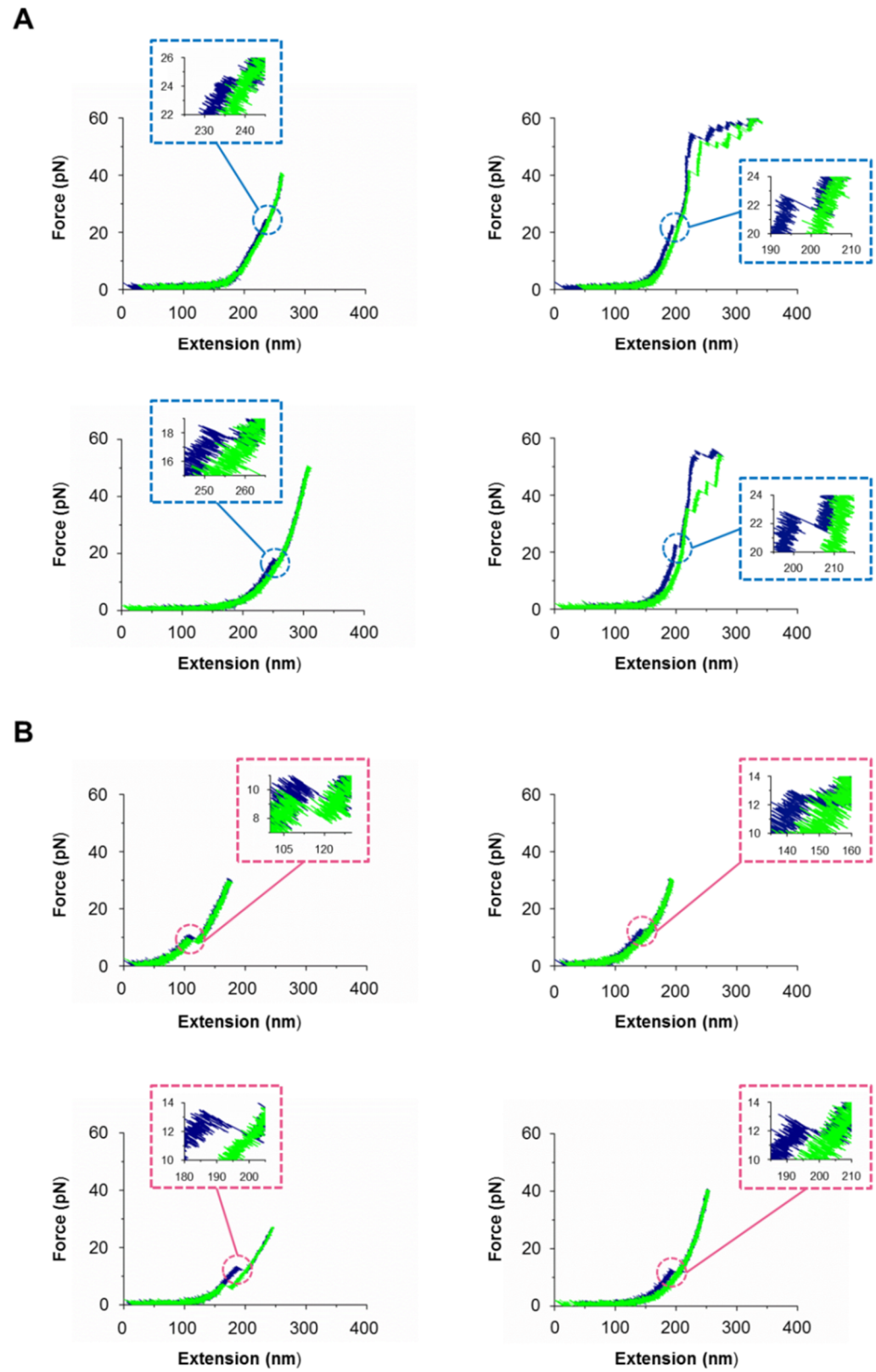

FIGURE 3.15. Replicates of force-extension curves of TERRA molecules with six GGGUUA repeats in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.

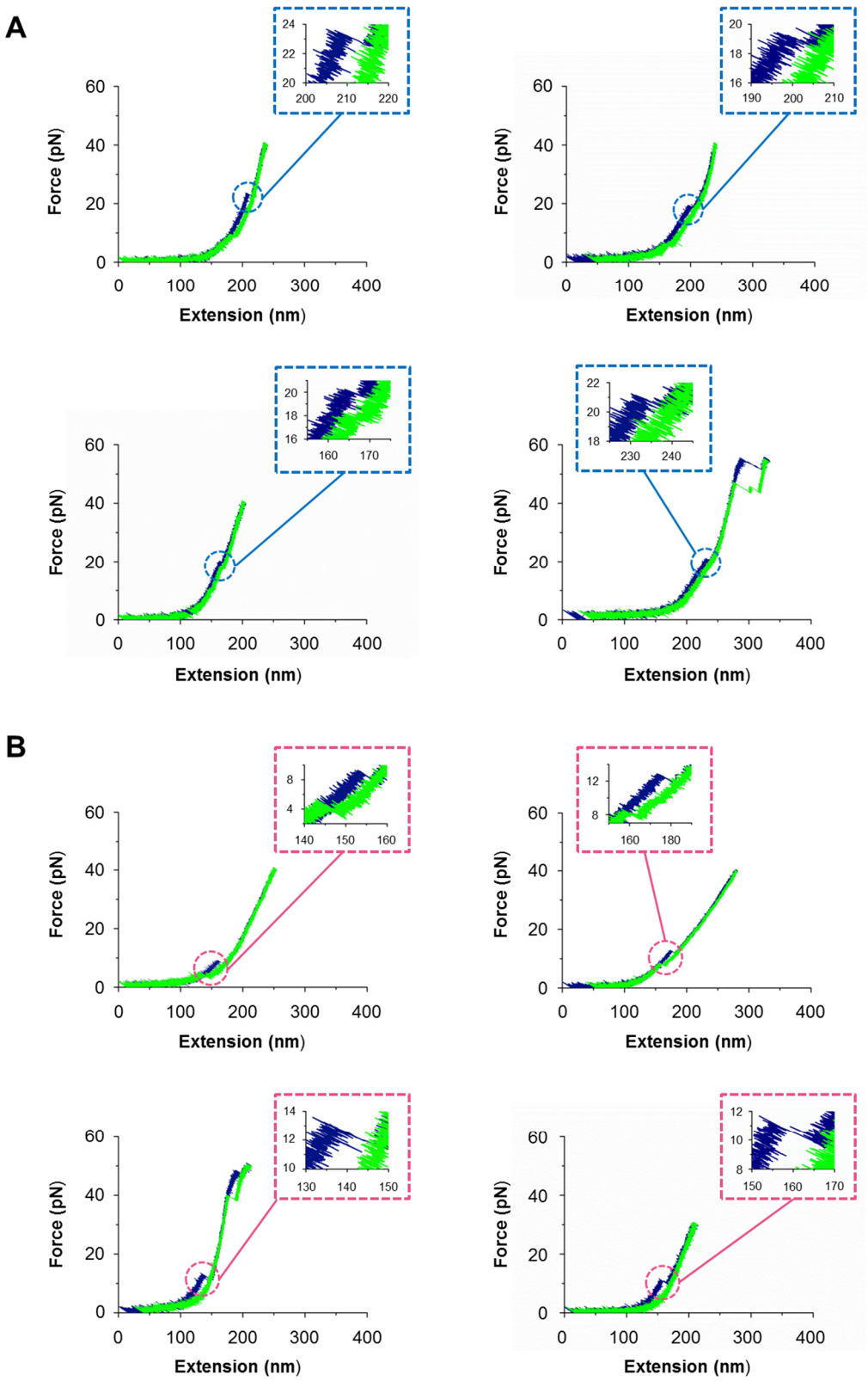

FIGURE 3.16. Replicates of force-extension curves of TERRA molecules with seven GGGUUA repeats in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.

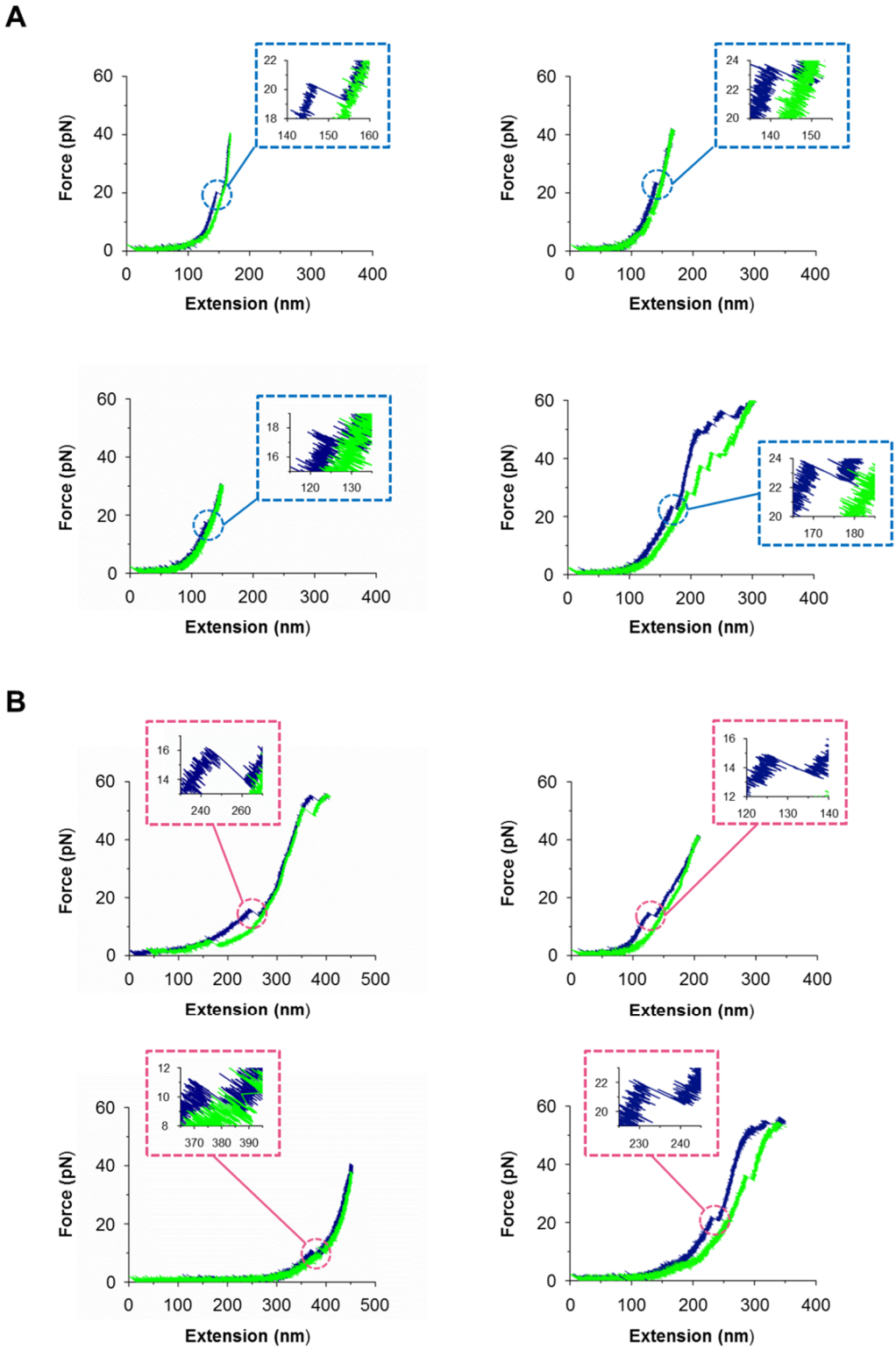

FIGURE 3.17. Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats and with more extra random ssRNA in which **(A)** a G-quadruplex/triplex or **(B)** single-stranded RNA condensed conformations are probably unfolded and refolded (blue and green lines, respectively). Each panel corresponds to a different molecule.

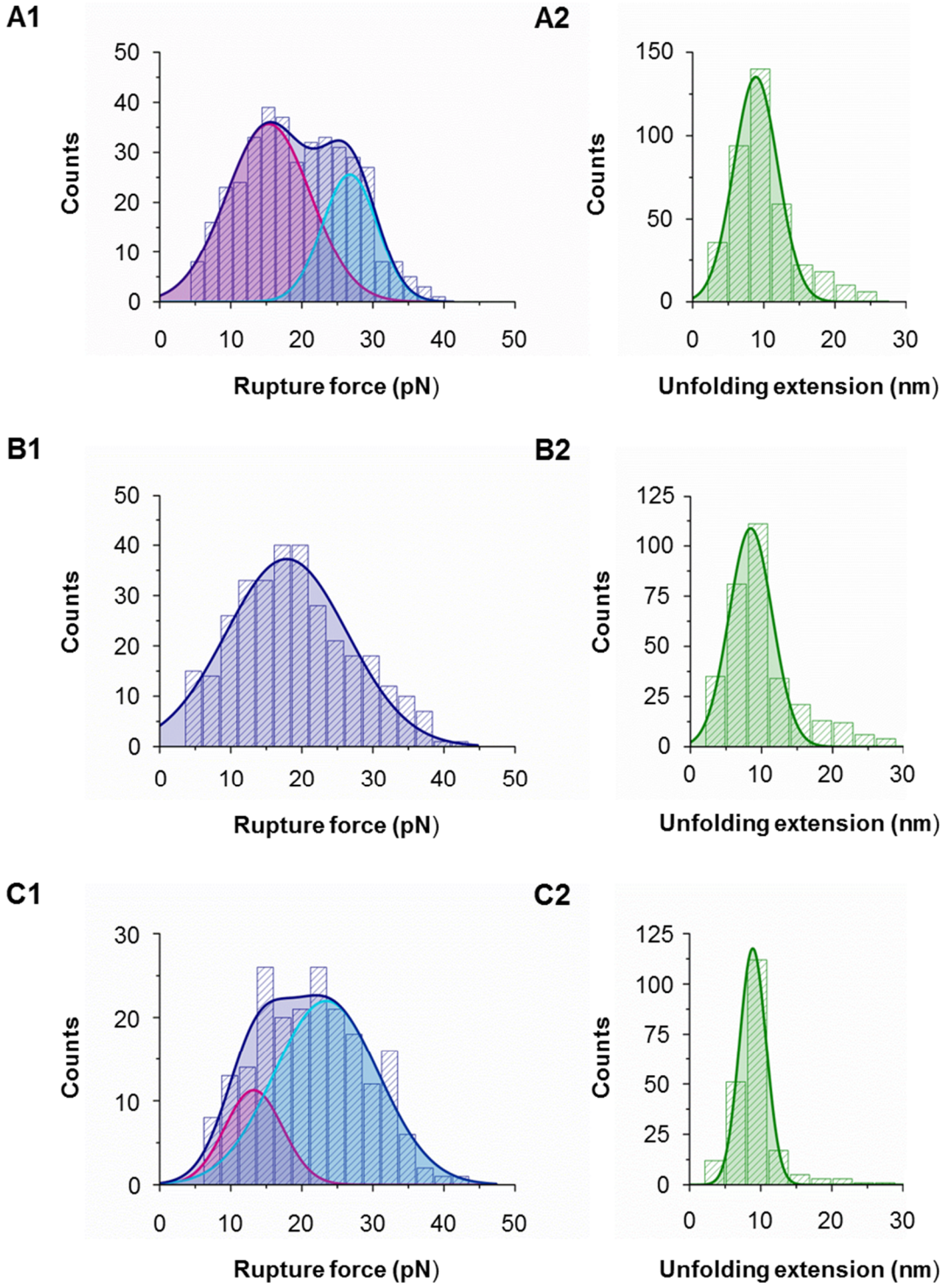

FIGURE 3.18. Statistics of single rupture events in RNA control systems. Rupture force (1) and unfolding extension distribution (2) for the RNA constructions of (A) 6 repeats (N=385), (B) 7 repeats (N= 317) and (C) 5 repeats with more random extra ssRNA (N= 205). Mean ± s.e. and s.d. according to Gaussian fitting are: (A1) 15.3 ± 1.0 and 5.9 pN (pink line), 26.8 ± 0.9 and 3.9 pN (light blue line); (A2) 8.9 ± 0.2 and 3.1 nml; (B1) 17.9 ± 0.5 and 8.5 pN; (B2) 8.5 ± 0.3 and 3.0 nm; (C1) 13.2 ± 1.9 and 4.1 pN (pink line), 23.5 ± 3.8 and 7.3 pN (light blue line) and (C2) 8.9 ± 0.1 and 1.9 nm. In (A1) and (C1) dark blue line corresponds to Gaussian double-peak fitting.

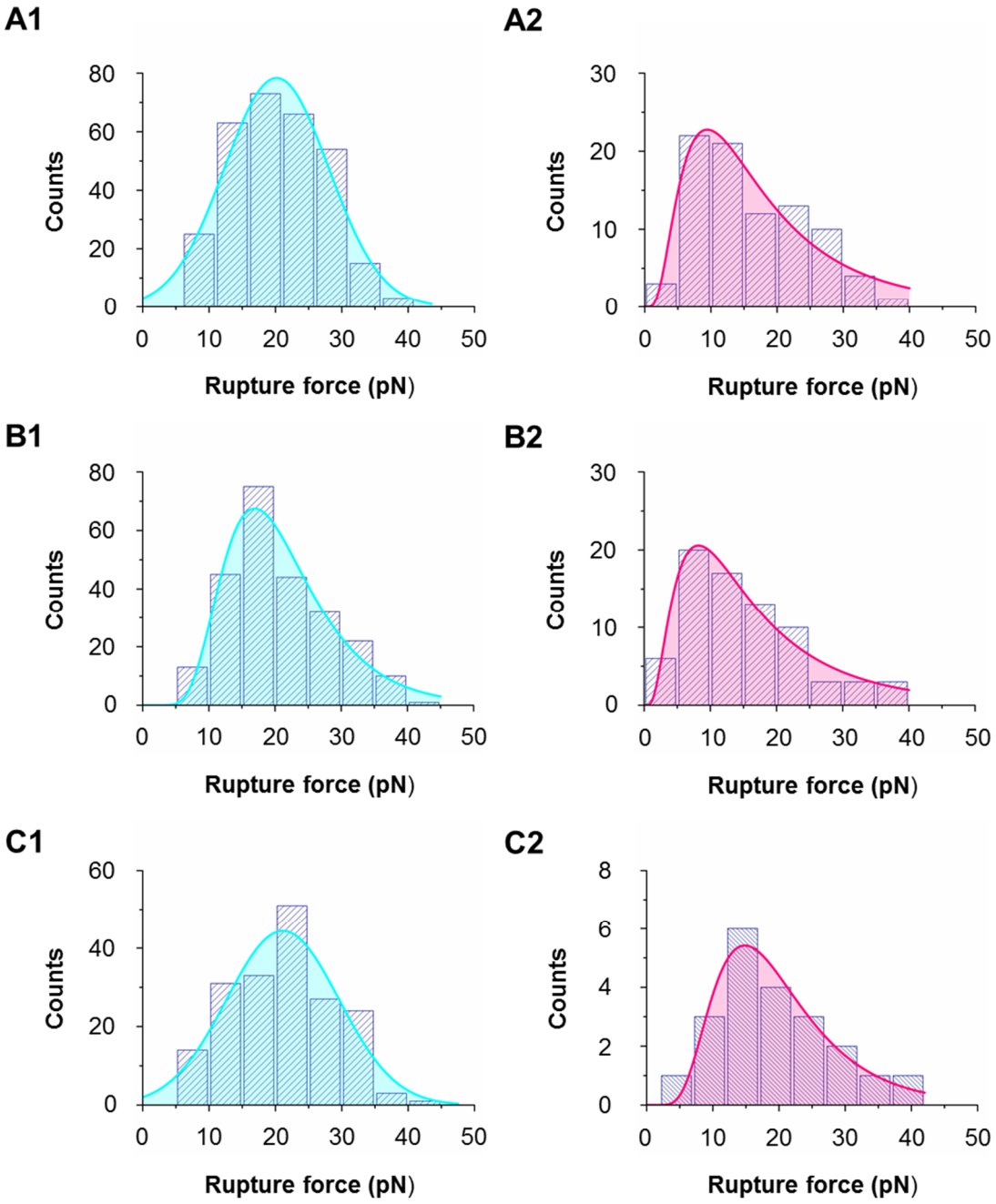

FIGURE 3.19. Rupture force distributions for the different conformations in RNA control systems. G-quadruplex and G-triplex formed by **(A1)** RNA molecules with six GGGUUA repeats (N= 299, 77.7%), **(B1)** RNA molecules with seven GGGUUA repeats (N= 242, 76.3%) and **(C1)** RNA molecules with five GGGUUA repeats and more extra random ssRNA (N= 184, 89.8%), data filtered in the range [0, 12] nm. Duplex or other condensed conformations formed by **(A2)** RNA molecules with six repeats (N= 86, 22.3 %), **(B2)** RNA molecules with seven repeats (N=75, 23.7%) and **(C2)** RNA molecules with five repeats and more extra random ssRNA (N= 21, 10.2%), data filtered in the range (12, 24] nm. Gaussian regressions have been used in A1 and C1 with mean and s.d.: (A1) 20.2 and 7.9 pN and (C1) 21.1 and 8.5 pN. Log-normal regressions have been applied to B1 and A2-C2 with median and s.d.: (A2) 15.1 and 14.7 pN, (B1) 19.7 and 8.7 pN, (B2) 14.0 and 15.3 pN, (C2) 18.3 and 9.9 pN. Peak values (mode) ± s.e. are: (A1) 20.2 ± 0.6 pN, (A2) 9.4±± 1.3 pN, (B1) 16.9 ± 0.6 pN, (B2) 8.2 ± 0.9 pN, (C1) 21.1 ± 0.9 pN and (C2) 14.8±± 0.9 pN.

To further confirm these results, *Figure 3.19* presents the histograms resulting from the application of the filter previously used for the reference molecule. The three sequences display separate single-peak distributions: a high force peak at around 20 pN, which as discussed, is attributed to the G-quadruplex/triplex, and a low force peak. The latter, increases with the number of nucleotides in the sequences: from ~9 pN for the reference sequence (*Fig. 3.12 (B)*), to ~15-18 pN for these longer molecules.

More importantly, from these histograms, we reveal that by increasing the number of repeats, the probability of stochastic blockage of G-quadruplex/triplex formation monotonically increases (19.4%, 22.3% and 23.7% for the 5, 6 and 7-repeat molecules, respectively). The addition of more random sequence, as also shown in *Figure 3.18 (C1)*, does not follow this trend. In particular, the 5-repeat molecule decreases the probability of stochastic G-quadruplex/triplex blockage from a 19.4% down to a 10.2% when 6 extra random nucleotides are added to the TERRA sequence. This is due to the fact that this molecule presents a lower number of Gs, which is a very promiscuous nucleobase thus tending to increase the probability of a condensate in RNA.

3.3.2. Consecutive rupture events

Figures 3.20-3.23 show characteristic force-extension curves corresponding to different molecules with sequential rupture events for the TERRA molecules with 5-7 repeats and extra random sequence.

As discussed above, the formation of condensed ssRNA or the hairpin represented in *Figure 3.12 (B)* (insets) blocks the assembly of the G-triplex and the G-quadruplex. Nevertheless, when one of these non-canonical structures assembles, it may interact with the excess ssRNA in our sequence. Then, the mechanical unfolding should provoke the separation of the extra ssRNA from the G-quadruplex or G-triplex before the full opening of the structure. According to the rupture forces and unfolding extensions obtained from force-extension curves as those represented in *Figures 3.20-3.23*, this is the most probable scenario for the molecules analysed. The statistical analysis of the 5-repeat TERRA molecule when two consecutive rupture events are observed in the force-extension curves, *Figure 3.24*, confirms this interpretation: the first rupture peak complies with low forces and low extensions, *Figure 3.24 (A)*, and the second one with high forces and typical extensions for the G-quadruplex and G-triplex, *Figure 3.24 (B)*.

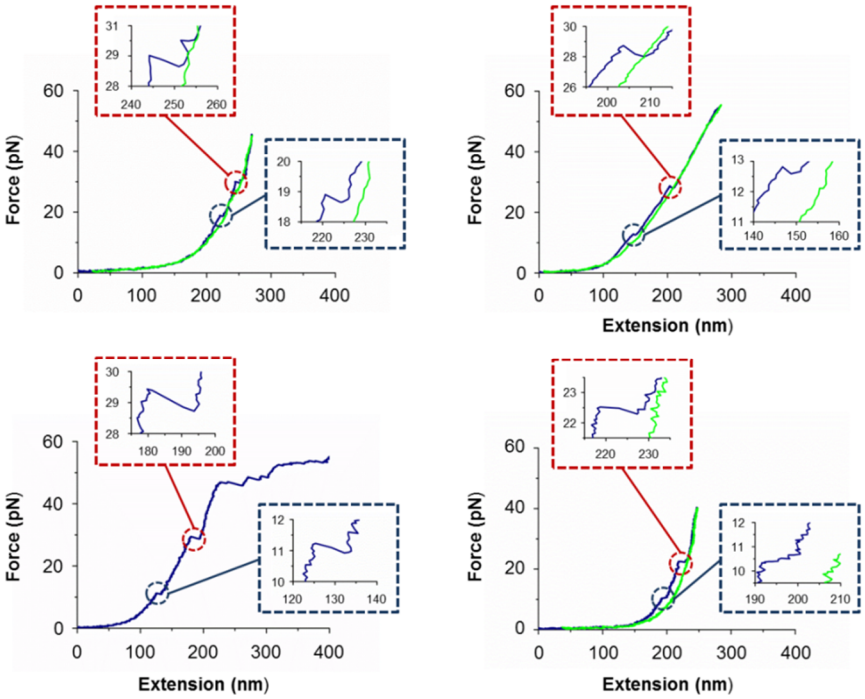

FIGURE 3.20. Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats in which sequential unfolding events take place. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.

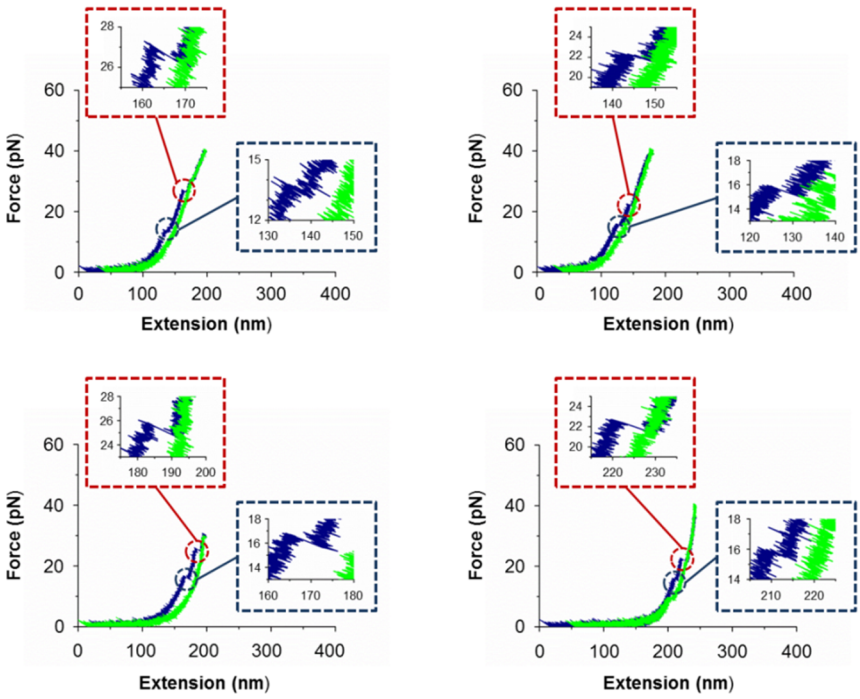

FIGURE 3.21. Replicates of force-extension curves of TERRA molecules with six GGGUUA repeats in which sequential unfolding events take place. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.

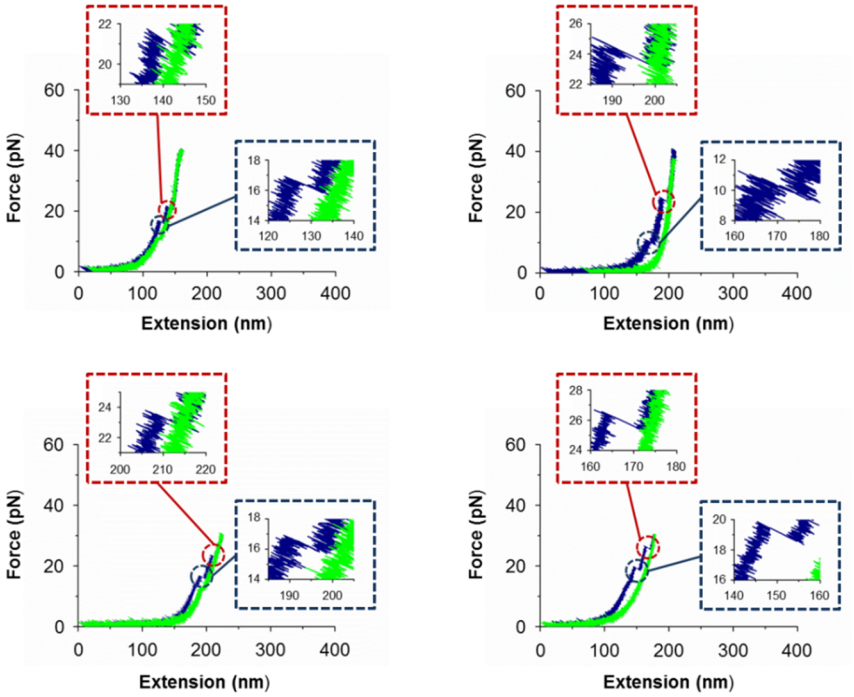

FIGURE 3.22. Replicates of force-extension curves of TERRA molecules with seven GGGUUA repeats in which sequential unfolding events take place. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.

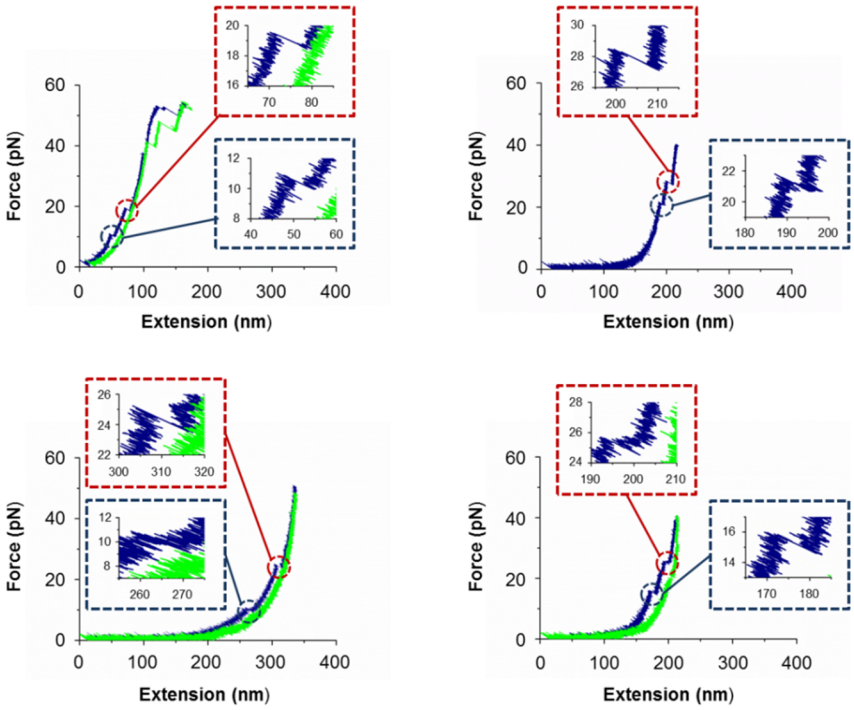

FIGURE 3.23. Replicates of force-extension curves of TERRA molecules with five GGGUUA repeats and more extra random ssRNA in which sequential unfolding events take place. Each panel corresponds to a different molecule. Blue traces correspond to stretching paths and green ones to relaxation paths.

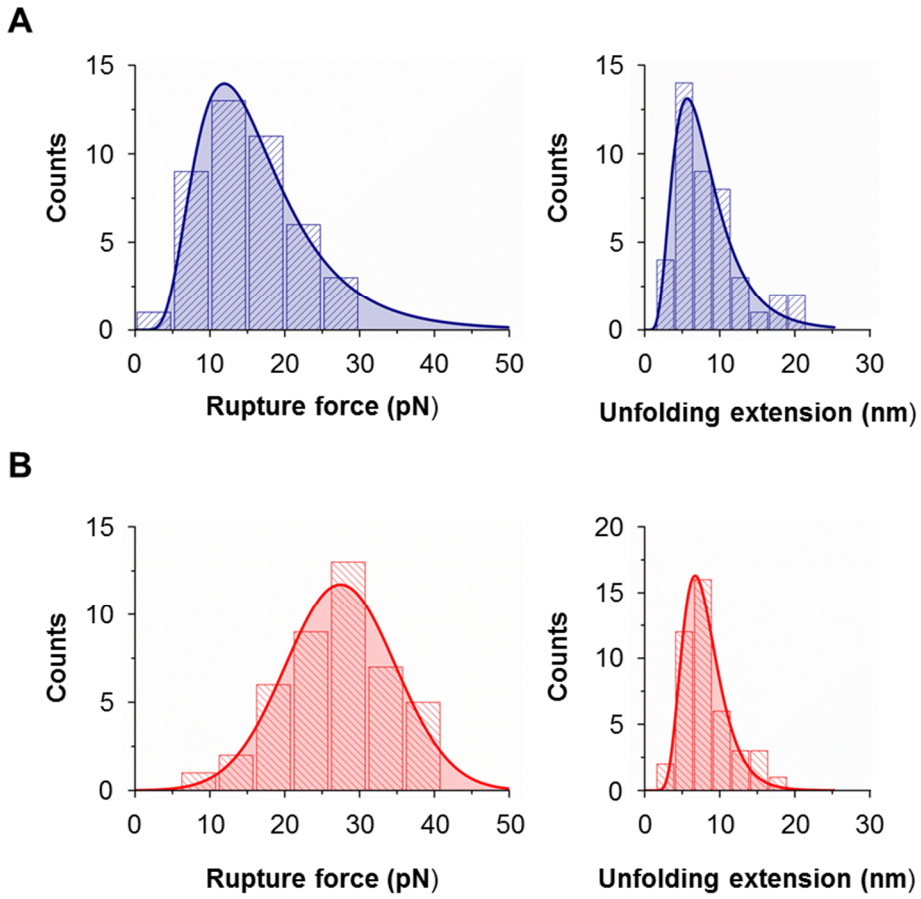

FIGURE 3.24. Statistics of single-molecule sequential unfolding events in the TERRA reference construct. (A) Rupture force and unfolding extension distributions for the first rupture event. Median and s.d. according to a log-normal fitting for this force distribution is 14.9 and 8.3 pN (peak value (mode) ± s.e., 11.9 ± 0.5 pN). Median and s.d. for the unfolding extension distribution is 7.3 and 4.4 nm (peak value (mode) ± s.e., 5.7 ± 0.4 nm). **(B)** Rupture force and unfolding extension distributions for the second rupture event. Mean ± s.e. and s.d. for this force distribution, according to a Gaussian fitting (red line) is 27.4 ± 0.5 and 7.3 pN. Median and s.d. according to log-normal fitting for the unfolding extension distribution are 7.6 and 2.8 nm (peak value (mode) ± s.e., 6.7 ± 0.2 nm).

The statistics corresponding to the 6 and 7-repeat TERRA molecules and that with 5 repeats and extra random sequence, *Figure 3.25*, confirm this trend. Importantly, it is observed that the probability of interaction of the formed G-quadruplex with the excess single-stranded increases with the number of added nucleotides: 12.6%, 17.4% and 26.1% for 5, 6 and 7 repeats, and 27% for the 5-repeat molecule with more extra random ssRNA, a trend which was expected for longer molecules since self-interactions allowed by the bending elasticity are more probable.

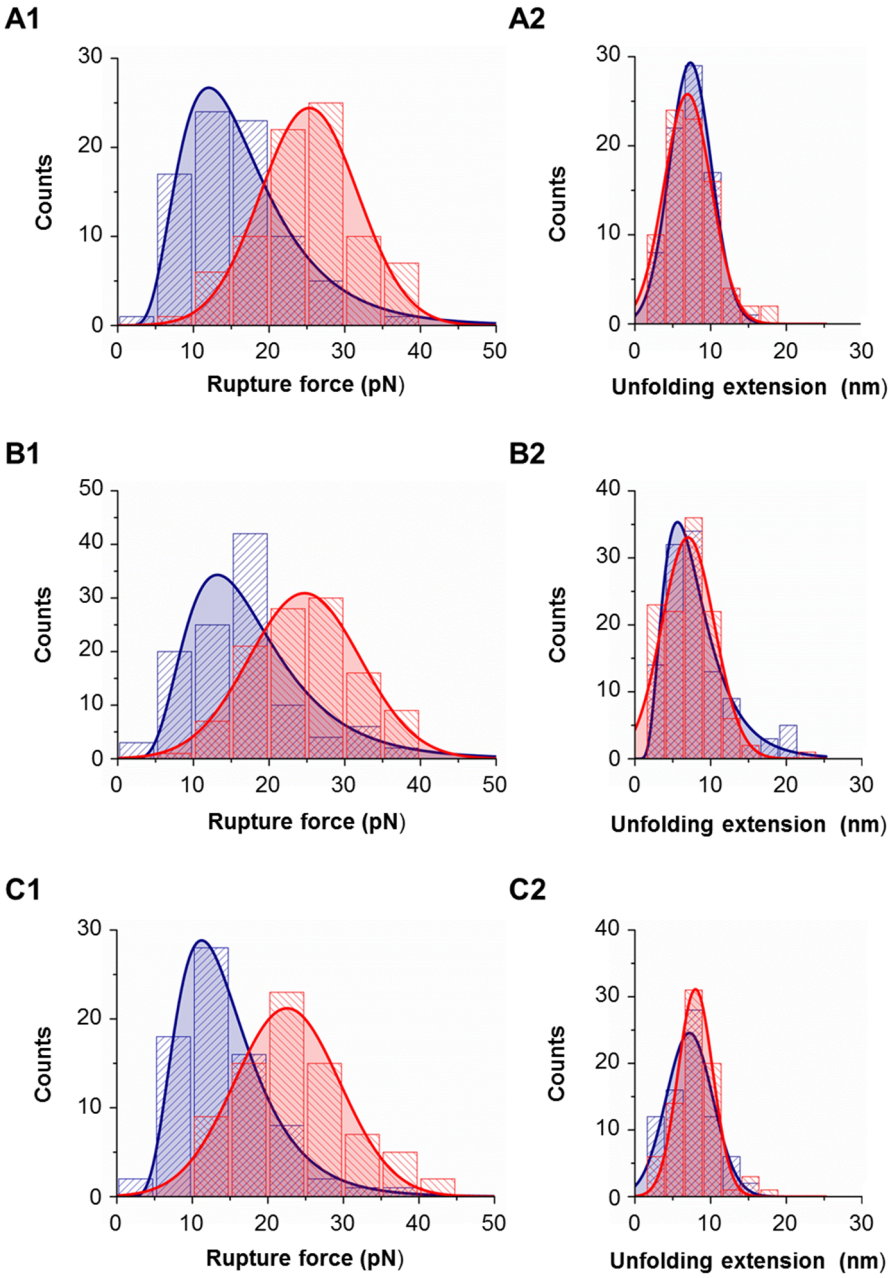

FIGURE 3.25. Statistics of single-molecule sequential unfolding events in control sequences. Rupture force distributions for RNA molecules with **(A1)** six GGGUUA repeats (N=81), **(B1)** seven GGGUUA repeats (N= 112) and **(C1)** five GGGUUA repeats and more extra random ssRNA (N=76). Median and s.d. for the first peak (blue) according to log-normal fittings are (A1) 14.9 and 8.2 pN, (B1) 16.1 and 8.4 pN and (C1) 13.4 and 6.3 pN. Peak values (mode) ± s.e. are: 12.0 ± 0.6 pN, 13.2 ±± 1.5 pN and 11.2 ±± 0.2 pN, respectively. Mean ± s.e. and s.d. for the second peak (red) according to a Gaussian fitting are (A1) 25.3 ± 0.5 and 6.4 pN, (B1) 24.7 ± 0.3 and 7.4 pN and (C1) 22.5 ±± 0.5 and 7.0 pN. Unfolding extension distributions for RNA molecules with **(A2)** six GGGUUA repeats, **(B2)** seven GGGUUA repeats and **(C2)** five GGGUUA repeats and more extra random ssRNA. Mean ± s.e. and s.d. for the first peak (blue) according to Gaussian fittings are: (A2) 7.30 ± 0.04 and 2.8 nm, (C2) 7.2 ± 0.3 and 3.1 nm; median and s.d. according to a log-normal fitting is (B2) 7.2 ± 4.2 nm (peak value (mode) ± s.e., 5.6 ± 0.4 nm). Mean ± s.e. and s.d. for the second peak (red) according to a Gaussian fitting are (A2) 6.9 ± 0.2 and 3.1 nm, (B2) 7.0 ± 0.4 and 3.5 nm and (C2) 8.0 ± 0.2 and 2.3 nm.

3.4. Discussion

A strong conformational competition between condensed ssRNA and non-canonical structures has been found in the folding pathways of short TERRA molecules. In particular, it has been demonstrated that if a ssRNA condensed form is established in a RNA segment with *n* repetitions of the sequence GGGUUA, $4 < n < 8$, it blocks the assembly of both the G-quadruplex and the G-triplex, even considering that these latter conformations are more stable than the former for equal number of nucleotides. The number of TERRA repeats studied here is representative of the stochastic blockage phenomena since a number equal or greater than eight increases the configurational probability space and thus, the complexity of the simultaneous configurations.

The probability of G-quadruplex/G-triplex formation, P(G-QT), can be calculated based on the next heuristic inference:

$$P(\text{G-QT}) = P(\text{consecutive RE}) + P(\text{single RE}) \times P(\text{G-QT}|\text{single RE}) \tag{2}$$

where *G-QT* stands for G-quadruplex or G-triplex, *RE* stands for rupture event and the vertical bar denotes conditional probability. The results, along with the previously presented experimental probabilities, are shown on *Table 3.5*. It is observed that the number of repeats hardly affects the probability of formation of the G-quadruplex/triplex, slightly above an 80%, and that the addition of random nucleotides flanking the hexanucleotide repeats attenuates the stochastic blockage, raising the probability of the non-canonical structures to a 90%.

	$(GGGUUA)_5$	$(GGGUUA)_6$	$(GGGUUA)_7$	$(GGGUUA)_5$ + ssRNA
P(Consecutive RE)	0.126	0.174	0.261	0.270
P(Single RE)	0.874	0.826	0.739	0.730
P(G-QT | Single RE)	0.806	0.777	0.763	0.898
P(G-QT)	**0.830**	**0.815**	**0.825**	**0.925**
P(random coil)	**0.170**	**0.185**	**0.175**	**0.075**

TABLE 3.5. Probability of G-quadruplex/triplex formation. P, probability; RE, rupture event; G-QT, G-quadruplex or G-triplex; "|", conditional probability.

Single rupture events as those represented in *Figures 3.9* and *3.15-3.17* may correspond to separate events sequential but proximal in force, which would match to the release of one arm of the G-quadruplex rapidly followed by the rupture of the G-triplex into ssRNA. Neither these events nor the G-hairpin unfolding [279-281] could be distinguished within our experimental

resolution. Although some force-extension curves with two rupture events in our experiments could be interpreted in this fashion, we have not found sufficient statistical significance. Then, we conclude that, if in-pathway, G-triplex and G-hairpin breakage takes place almost simultaneously after the G-quadruplex disassembly in our experiments; that is, the sequential release of the GGGUUA arms of the G-quadruplex are detected within the same rupture event in the force-extension curves at the herein used experimental velocity [287]. A triangular, three state kinetic model, as that proposed elsewhere [284, 309, 310, 316], is compatible with our results (*Fig.3.26*).

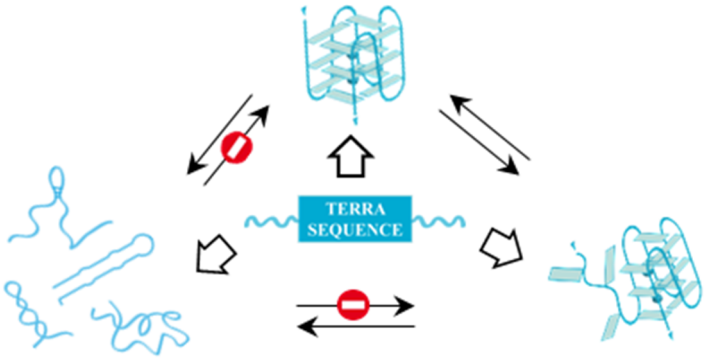

Figure 3.26. Schematic representation of the triangular, three state kinetic model of the mechanical unfolding of TERRA molecules.

The dynamic behaviour of the rupture events has another feature in the stretch-relaxation cycles: the reversibility. Refolding during relaxation is the normal behaviour but the mechanical hysteresis is typically lower for the duplexes with respect to the G-quadruplex or G-triplex conformations, as measured by the degree of overlapping of the stretch and relaxation traces (see *Fig. 3.9* and *3.15-3.17*). The refolding is more likely for the duplex than for the G-triplex or G-quadruplex since these latter conformations involve the ordered assembly of three and four strands, respectively, within similar time frames, which is less probable than the assembly of randomly-condensed ssRNA or duplexes, with lower entropy. Then, it is expected that the non-canonical conformations involve refolding further from equilibrium than simpler conformations at the same loading rate.

In these conditions, the free energy of unfolding of each species can be recovered by using ***Jarzynski's equality*** (2), which evaluates this equilibrium thermodynamic potential increment from non-equilibrium work measurements as follows [317]:

$$\Delta G_{unfold} = -k_B T \ln \sum_{i=1}^{N} \frac{1}{N} \exp\left(-\frac{W_i}{k_B T}\right) \quad (3)$$

where ΔG_{unfold} is the free-energy difference between the folded and unfolded states of the corresponding conformation; k_BT, the product of the Boltzmann's constant and the temperature (0.592 kcal/mol at 298 K); *N,* the number of unfolding data compiled for the conformation under study and *W,* the amount of non-equilibrium work done to unfold the structure (G-quadruplex/triplex or RNA random coils). The latter was calculated from the force and extension signatures in the stretching paths of the force-extension curves for the different conformations in each molecular preparation. *Tables 3.6* and *3.7* collect these data for the non-canonical structure and the condensates, respectively. *Tables 3.8* and *3.9* provide the resulting Gibbs free energy increments. Right columns show energies per nucleotide.

	W [kcal/mol] (mean ± s.e.)	**W[kcal/(mol·nt)] (mean ± s.e.)**
5 RNA telomeric repeats (n=241)	25.7 ± 0.8	1.22 ± 0.04
5 DNA telomeric repeats (n=163)	28.4 ± 1.0	1.35 ± 0.05
5 RNA telomeric repeats + random ssRNA (n=184)	25.7 ± 0.8	1.22 ± 0.04
6 RNA telomeric repeats (n=299)	23.4 ± 0.6	1.12 ± 0.03
7 RNA telomeric repeats (n=242)	22.5 ± 0.7	1.07 ± 0.03

TABLE 3.6. Non equilibrium work done to unfold the G-quadruplex structure (s.e., standard error).

	W [kcal/mol] (mean ± s.e.)	**W[kcal/(mol·nt)] (mean ± s.e.)**
5 RNA telomeric repeats (n=241)	37.6 ± 3.2	0.99 ± 0.08
5 RNA telomeric repeats + random ssRNA (n=184)	46.5 ± 4.4	1.06 ± 0.10
6 RNA telomeric repeats (n=299)	37.1 ± 2.0	0.84 ± 0.05
7 RNA telomeric repeats (n=242)	35.8 ± 2.2	0.72 ± 0.04

TABLE 3.7. Non equilibrium work done to unfold condensed structures.

	ΔG_{unfold} [kcal/mol] (mean ± s.e.)	**ΔG_{unfold} [kcal/(mol·nt)] (mean ± s.e.)**
5 RNA telomeric repeats (n=241)	7.5 ± 0.4	0.36 ± 0.02
5 DNA telomeric repeats (n=163)	6.9 ± 0.6	0.33 ± 0.03
5 RNA telomeric repeats + random ssRNA (n=184)	8.1 ± 0.3	0.39 ± 0.02
6 RNA telomeric repeats (n=299)	9.3 ± 0.3	0.44 ± 0.01
7 RNA telomeric repeats (n=242)	7.4 ± 0.4	0.35 ± 0.02

TABLE 3.8. Change in free energy of G-quadruplex structure calculated by the Jarzinsky's equation.

	ΔG_{unfold} [kcal/mol] (mean ± s.e.)	**ΔG_{unfold} [kcal/(mol·nt)] (mean ± s.e.)**
5 RNA telomeric repeats (n=58)	12.3 ± 0.6	0.32 ± 0.02
5 RNA telomeric repeats + random ssRNA (n=21)	12.6 ± 0.6	0.29 ± 0.01
6 RNA telomeric repeats (n=86)	10.8 ± 0.6	0.25 ± 0.01
7 RNA telomeric repeats (n=75)	11.1 ± 0.6	0.22 ± 0.01

TABLE 3.9. Change in free energy of random coil structures calculated by the Jarzinsky's equation.

Results for the DNA analogue 5-repeat molecule yields $\Delta G_{unfold} = -(6.9 \pm 0.6)$ kcal/mol, which is similar within experimental error with earlier bulk reports that yielded −7.1 kcal/mol for the parallel G-quadruplex assembled with a four-repeat molecule [318]. This consistency serves as a control check, both from a qualitative and a quantitave viewpoint, for our results at the single-molecule level. RNA G-quadruplex has always higher Gibbs free energy than the DNA analogue, also consistent with earlier literature [284].

Conclusions

TElomeric Repeat containing RNA (TERRA) molecules have been studied by single-molecule microscopic and force-spectroscopic techniques. The interest in these non-coding transcripts with subtelomere-derived sequence lies in their potential biological roles, such as telomeric heterochromatin formation and telomere protection, which are probably associated to their capacity to fold into multiple G-quadruplex blocks. Therefore, to better understand these RNA species in terms of structure and mechano-chemical stability, their analysis by atomic force microscopy (AFM) and optical tweezers (OT) has been herein performed.

According to the results obtained in this thesis, it can be concluded that:

1. In contrast to the expected scenario in which the TERRA sequence folds into a parallel G-quadruplex or a G-triplex, a remarkable conformational competition between unspecific self-interacting ssRNA species and these non-canonical conformations associated with the TERRA repeats, $(GGGUUA)_n$, has been observed for the first time and at the single-molecule level.

2. This interference represents a stochastic blockage in the G-quadruplex folding pathways of up to 20% due to the condensation of TERRA single strands. The formation of a condensed ssRNA 'knot' instead of a G-quadruplex/triplex can largely facilitate its denaturation by specialized protein machinery; specifically, decondensing the ssRNA involves an average force barrier near 10 pN, whereas that of the G-triplex or the G-quadruplex is on average above 20 pN.

3. This behaviour is unique to the RNA G-quadruplex species since we have not found an analogue behaviour in their DNA counterparts herein synthesised.

4. Application of Jarzynski's equality reveals that the TERRA G-quadruplex free-energy per nucleotide is larger than that of the condensates. A good agreement has been found with earlier literature in the use of this analysis to DNA analogue

molecules, in which equilibrium free energies were obtained by ensemble-averaged experiments, thus supporting the validity of our mechanical stability analysis on the TERRA molecules. In tune with the profile of the stretch-relaxation cycles, this analysis also confirms that our single-molecule force-induced unfolding protocol takes place away from equilibrium.

5. G-quadruplex unfolding is a halfway problem in complexity between protein and duplex nucleic acid unfolding.

6. With regards to AFM results, a strategy to analyse RNA G-quadruplex/triplex structures on mica surface at the single-molecule level has been determined. We believe our methodology is useful not only to cross-check G-quadruplex molecular preparations but also for applications in nanotechnology that make use of non-canonical conformations as building blocks over a surface under controlled conditions.

Conclusiones

En esta tesis se han estudiado moléculas de ARN con diversas repeticiones de la secuencia telomérica GGGUUA, comúnmente denominadas *TERRA* (del inglés, *TElomeric Repeat-containing RNA*), mediante la aplicación de técnicas de moléculas individuales, tanto microscópicas como de espectroscopia de fuerza. El interés en estos ARN no codificantes reside en sus potenciales funciones biológicas, tales como la formación de heterocromatina en los telómeros o la protección de estos últimos, las cuales probablemente estén asociadas con la capacidad de plegamiento de dichos ARNs en tándems de G-quadruplexes. Por tanto, para un mayor entendimiento de estas moléculas en términos de estructura y estabilidad mecano-química, se ha procedido a su análisis mediante microscopía de fuerza atómicas (AFM; del inglés, *Atomic Force Microscopy*) y pinzas ópticas (OT; del inglés, *Optical Tweezers*).

De acuerdo a los resultados obtenidos en esta tesis, se concluye que:

1. En contra del patrón de plegamiento esperado para una secuencia *TERRA*; esto es, bien en forma de *G-quadruplex* paralelo o como *G-triplex*, se ha observado, por primera vez y a nivel de molécula individual, una notable competencia conformacional entre dichas estructuras no canónicas y aquellas resultantes de interacción inespecífica inherente a hebras sencillas de ARN.

2. Esta interferencia se traduce en un bloqueo estocástico del plegamiento de las moléculas *TERRA* en forma de estructuras tipo *G-quadruplex/triplex* de hasta un 20%. La formación de un condensado de ARN, en lugar de un *G-quadruplex/triplex*, podría facilitar, en gran medida, su desnaturalización mediante complejos proteicos específicos, ya su barrera de fuerza se encuentra próxima a los 10 pN frente a los más de 20 pN del *G-triplex* o del *G-quadruplex*.

3. Este comportamiento es inherente a la naturaleza ribonucleica de las moléculas estudiadas, pues no se ha encontrado el mismo patrón en sus análogos químicos de ADN.

4. Mediante la aplicación del teorema de Jarzynski, se ha obtenido que la energía libre por nucleótido de un *G-quadruplex* de *TERRA* es mayor que la de los condensados de ARN. Paralelamente, se ha encontrado un gran nivel de concordancia entre la energía libre del sistema análogo de ADN resultante de aplicar dicho teorema y los datos existentes en bibliografía procedentes de experimentos de espectroscopía visible. Esto respalda la validez de nuestro análisis de estabilidad mecánica en moléculas *TERRA*. Por último, este análisis termodinámico también confirma que nuestro protocolo de despliegue inducido por la aplicación de fuerza a una sola molécula tiene lugar fuera del equilibrio, lo cual se refleja experimentalmente en la histéresis de las curvas fuerza-extensión.

5. Con respecto a los resultados de AFM, se ha desarrollado una estrategia exitosa para comprobar la formación de estructuras tipo *G-quadruplex/triplex* en moléculas *TERRA*. Esta metodología podría ser útil también para aplicaciones en nanotecnología, en las que dichas estructuras serían usadas para la construcción de sistemas modulares bajo condiciones controladas.

Appendix A

Mesoscopic model for DNA GQ unfolding

Currently, single-molecule characterization is limited by the temporal and spatial resolutions of the experimental techniques. Because of that, resorting to theoretical modelling to obtain complementary information about the system under study is necessary. Focusing on GQs, they have been modelled at different temporal and length scales, from quantum calculations [319, 320] and molecular dynamics simulations [321-326] to mesoscopic approaches [327, 328].

From the force-induced unfolding experiments obtained in this thesis on a DNA telomeric sequence capable of forming a GQ (see Chapter 3), a mesoscopic model to describe both its mechanical and thermal stability has been developed by the group led by Dr. A. Fiasconaro and Dr. F. Falo at Zaragoza University [329]. In addition to our experimental results, model parameters were obtained taking into account a previous atomistic study of telomeric DNA GQ performed in this physical-theoretical research group [326] and the melting temperatures of DNA GQ [330]. As shown in *Fig.A.1*, the mesoscopic model is characterized by single beads representing each nucleotide of the ssDNA chain and the monovalent cations located at the central channel of the GQ structure.

Being *bead 1* and *bead 21*, the first and last nucleotides in the telomeric DNA chain (see *Fig. A.1*), pulling simulations were performed by fixing the latter to the origin of coordinates, whereas the former is attached to a harmonic spring k_A whose end is displaced at constant velocity, υ (see *Fig.A.2*). The component of the force along the pulling direction was calculated as:

$$F = k_A \cdot \left(x_A(t) - x_1(t)\right) \tag{A.1}$$

where $x_A(t)$ and $x_1(t)$ are the components of the distance along the pulling direction of the spring end (point A) and the pulled bead 1, respectively, as depicted in *Figure A.2*.

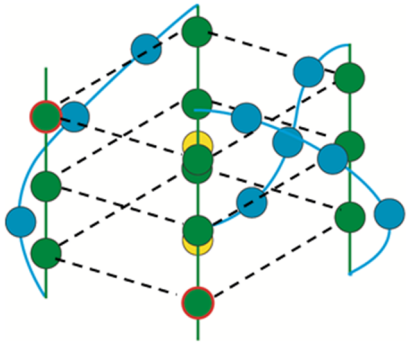

Figure A.1. Scheme of a parallel GQ assembly. Each nucleotide is represented by a single bead (not to scale). Green beads correspond to guanines and the blue ones, to the nucleotides that form the loops; i.e. TTA Yellow beads symbolized monovalent cations. Beads 1 and 21 have been highlighted in red because of their involvement in the pulling simulations (see the text for details). For simplicity, the twist between successive G-tetrad planes is not represented.

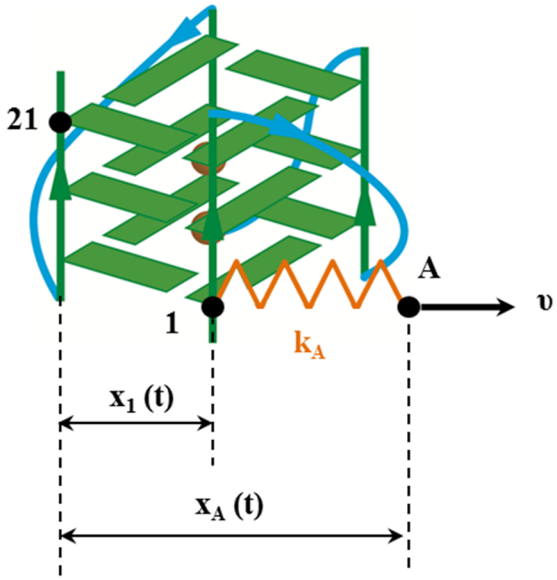

Figure A.2. Scheme of the pulling simulations. The first and last nucleotides in the telomeric DNA chain are represented by black beads and labelled as 1 and 21, respectively. Bead 21 is fixed to the origin of coordinate, whereas bead 1 is attached to a harmonic spring k_A whose end A is displaced at a constant velocity υ. The distance projections along the pulling direction, x_A and x_1, are calculated during the simulations.

The elastic constant, k_A, was set to 0.4 pN/nm, in accordance to the experimental value estimated from the slope of the force-extension curves in the enthalpic elasticity regime before

the rupture event. *Figure A.3* shows the comparison between two simulations at υ=0.08 and the experimental data, showing a clear agreement on the values of the unfolding force.

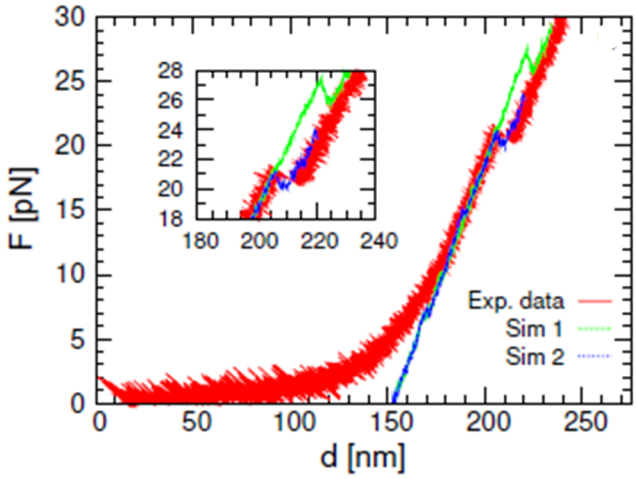

Figure A.3. Comparison of the force-extension curve between two simulations at υ=0.08 and one experimental realization of the mechanical unfolding. The experimental curve (in red) shows superimposed the entropic elasticity regime of the double-stranded handles at low forces.

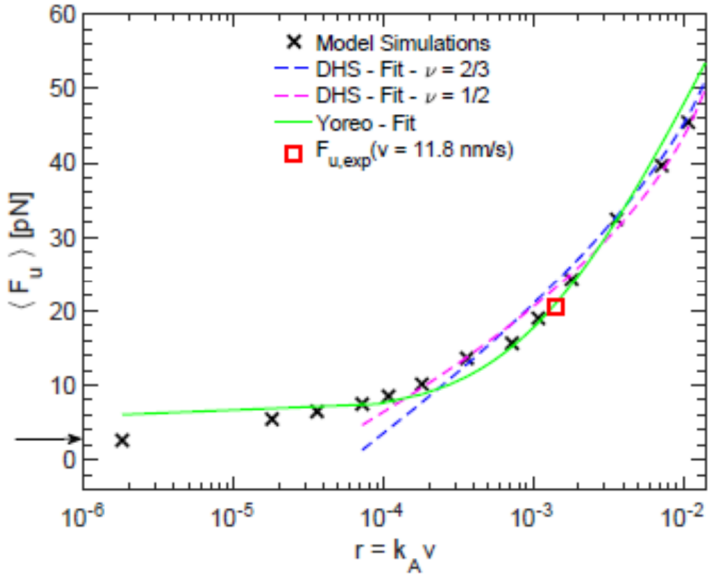

Figure A.4. Mean value of the unfolding force, F_u, as a function of the pulling rate, r and fitting to the kinetic models Dudko-Hummer-Szabo (DHS) and Yoreo. Note that the four lowest values of the loading rates are excluded from the DHS fittings because in this region some rebinding events are observed during the unfolding, and, then, the kinetics models are not valid. Two experimental force values are included in this figure: F_u = 20.4 pN from the unfolding experiments carried out in this thesis (red square) and F_u = 2.5 pN from the constant force experiment of Long et al. [331] (arrow in the left axis). The loading rate *r* is expressed in dimensionless units. That is because its value at the experimental unfolding force of 20.4 pN is used to define the time units of the mesoscopic model.

To express the velocity in real units, it was necessary to specify the time units. For that purpose, the mean value of the unfolding force, F_u, was considered as a function of the velocity, or equivalently, as a function of the pulling rate $r = k_A \cdot \upsilon$ as shown in *Figure A.4*. The evaluated F_u as a function of r nicely reproduces the nonlinear increasing behaviour observed in dynamic force spectroscopy experiments at high velocities and the almost force independent regime at low velocities, behaviours that are well fitted by the *Dudko-Hummer-Szabo* (DHS) and the *Yoreo* models. In this figure, two experimental force values are also included: the unfolding force F_u = 20.4 pN obtained from the DNA experiments carried out in this thesis (see Chapter 3) at a pulling velocity υ = 11.8 nm/s —calculated by applying a linear fitting to extension vs time data near the rupture event— and the equilibrium force F_u = 2.5 pN obtained in the constant force experiment of Long et al. [331]. The loading rate in the dimensionless units of the mesoscopic model corresponding to this force value is r = 0.0014, and the velocity is $\upsilon = r/k_A$ = 0.0788 $l_u/t_u \approx 0.051$ nm$/t_u$. Equalling this value to the experimental one υ = 11.8 nm/s, they get the time unit to t_u= 0.0043 s.

Therefore, the mesoscopic model proposed here is able to reproduce unfolding forces in the same order of magnitude as in experimental studies, which are impossible with atomistic calculations [326]. In particular, it is possible to explore wide time scales and study the unfolding pathways by using loading rates up to five orders of magnitude lower than those allowed in microscopic simulations. Because of that, this theoretical model could be used as a tool for performing systematic studies on mechanical stability of different GQ conformations, either GQ geometries according to loop orientation (parallel vs antiparallel) or GQ tandems repeats [221], like those existing at the end of human telomeres.

Appendix B

List of publications

- Gutiérrez, I; Garavís, M; de Lorenzo, S.; Villasante, A. and Arias-Gonzalez, J.R. *Single-stranded condensation stochastically blocks G-quadruplex assembly in human telomeric RNA*. 2017. Pre-print.

- Bergues-Pupo, A.E; Gutiérrez, I.; Arias-Gonzalez, J.R.; Falo, F. and Fiasconaro, A. *Mesoscopic model for DNA G-quadruplex unfolding*. 2017. Pre-print.

Bibliography

1. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans*. nature, 1998. **391**(6669): p. 806-811.
2. Montgomery, M.K., S. Xu, and A. Fire, *RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans.* Proceedings of the National Academy of Sciences, 1998. **95**(26): p. 15502-15507.
3. Timmons, L. and A. Fire, *Specific interference by ingested dsRNA.* Nature, 1998. **395**(6705): p. 854-854.
4. Um, S.H., et al., *Enzyme-catalysed assembly of DNA hydrogel.* Nature materials, 2006. **5**(10): p. 797-801.
5. Park, N., et al., *A cell-free protein-producing gel.* Nature materials, 2009. **8**(5): p. 432-437.
6. Chhabra, R., et al., *DNA self-assembly for nanomedicine.* Advanced drug delivery reviews, 2010. **62**(6): p. 617-625.
7. Lee, J.B., et al., *DNA-based nanostructures for molecular sensing.* Nanoscale, 2010. **2**(2): p. 188-197.
8. Seeman, N.C., *Nanomaterials based on DNA.* Annual review of biochemistry, 2010. **79**: p. 65-87.
9. Roh, Y.H., et al., *DNAsomes: multifunctional DNA-based nanocarriers.* Small, 2011. **7**(1): p. 74-78.
10. Peng, S., et al., *From cells to DNA materials.* Materials Today, 2012. **15**(5): p. 190-194.
11. Church, G.M., Y. Gao, and S. Kosuri, *Next-generation digital information storage in DNA.* Science, 2012. **337**(6102): p. 1628-1628.
12. Goldman, N., et al., *Towards practical, high-capacity, low-maintenance information storage in synthesized DNA.* Nature, 2013. **494**(7435): p. 77-80.
13. Guo, P., *The emerging field of RNA nanotechnology.* Nature nanotechnology, 2010. **5**(12): p. 833-842.
14. Avery, O.T., C.M. MacLeod, and M. McCarty, *Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation*

by a desoxyribonucleic acid fraction isolated from pneumococcus type III. The Journal of experimental medicine, 1944. **79**(2): p. 137.

15. Hershey, A.D. and M. Chase, *Independent functions of viral protein and nucleic acid in growth of bacteriophage.* The Journal of general physiology, 1952. **36**(1): p. 39-56.
16. Zamenhof, S., G. Brawerman, and E. Chargaff, *On the desoxypentose nucleic acids from several microorganisms.* Biochimica et biophysica acta, 1952. **9**: p. 402-405.
17. Wilkins, M. and J. Randall, *Crystallinity in sperm heads: molecular structure of nucleoprotein in vivo.* Biochimica et biophysica acta, 1953. **10**: p. 192-193.
18. Watson, J. and F. Crick, *A structure for ADN.* Nature, 1953. **171**(4356): p. 737-8.
19. Wang, J.C., *Helical repeat of DNA in solution.* Proceedings of the National Academy of Sciences, 1979. **76**(1): p. 200-203.
20. Franklin, R.E. and R.G. Gosling, *The structure of sodium thymonucleate fibres. I. The influence of water content.* Acta Crystallographica, 1953. **6**(8-9): p. 673-677.
21. Franklin, R.E. and R.G. Gosling, *The structure of sodium thymonucleate fibres. II. The cylindrically symmetrical Patterson function.* Acta Crystallographica, 1953. **6**(8-9): p. 678-685.
22. Franklin, R.E. and R.G. Gosling, *Molecular configuration in sodium thymonucleate.* Nature, 1953. **171**(4356): p. 740-741.
23. Wahl, M.C. and M. Sundaralingam, *Crystal structures of A-DNA duplexes.* Biopolymers, 1997. **44**(1): p. 45-63.
24. Nekludova, L. and C.O. Pabo, *Distinctive DNA conformation with enlarged major groove is found in Zn-finger-DNA and other protein-DNA complexes.* Proceedings of the National Academy of Sciences, 1994. **91**(15): p. 6948-6952.
25. Timsit, Y., *DNA structure and polymerase fidelity.* Journal of molecular biology, 1999. **293**(4): p. 835-853.
26. Lu, X.-J., Z. Shakked, and W.K. Olson, *A-form conformational motifs in ligand-bound DNA structures.* Journal of molecular biology, 2000. **300**(4): p. 819-840.
27. Hormeno, S., et al., *Condensation prevails over BA transition in the structure of DNA at low humidity.* Biophysical journal, 2011. **100**(8): p. 2006-2015.
28. Hormeno, S., et al., *Mechanical Properties of High-G· C Content DNA with A-Type Base-Stacking.* Biophysical journal, 2011. **100**(8): p. 1996-2005.
29. Hormeño, S., et al., *Mechanical stability of low-humidity single DNA molecules.* Biopolymers, 2012. **97**(4): p. 199-208.
30. Arias-Gonzalez, J.R., *Single-molecule portrait of DNA and RNA double helices.* Integrative Biology, 2014. **6**(10): p. 904-925.

31. Wang, A., et al., *Molecular structure of a left-handed double helical DNA fragment at atomic resolution.* Nature, 1979. **282**(5740): p. 680.
32. Hardwick, J.S., et al., *5-Formylcytosine does not change the global structure of DNA.* Nat Struct Mol Biol, 2017. **advance online publication**.
33. Drew, H.R., et al., *Structure of a B-DNA dodecamer: conformation and dynamics.* Proceedings of the National Academy of Sciences, 1981. **78**(4): p. 2179-2183.
34. https://www.liverpool.ac.uk/~bates/Z-DNAframe.html.
35. Bikard, D., et al., *Folded DNA in action: hairpin formation and biological functions in prokaryotes.* Microbiology and Molecular Biology Reviews, 2010. **74**(4): p. 570-588.
36. Brázda, V., et al., *Cruciform structures are a common DNA feature important for regulating biological processes.* BMC molecular biology, 2011. **12**(1): p. 33.
37. Wells, R., et al., *The chemistry and biology of unusual DNA structures adopted by oligopurine. oligopyrimidine sequences.* The FASEB Journal, 1988. **2**(14): p. 2939-2949.
38. Htun, H. and J.E. Dahlberg, *Topology and formation of triple-stranded H-DNA.* Science, 1989. **243**(4898): p. 1571-1577.
39. Frank-Kamenetskii, M.D. and S.M. Mirkin, *Triplex DNA structures.* Annual review of biochemistry, 1995. **64**(1): p. 65-95.
40. Soyfer, V.N. and V.N. Potaman, *Triplex Recognition*, in *Triple-helical Nucleic Acids.* 1996, Springer. p. 151-193.
41. Gehring, K., J.-L. Leroy, and M. GuÃ, *A tetrameric DNA structure with protonated cytosine. cytosine base pairs.* Nature, 1993. **363**(6429): p. 561.
42. Leroy, J.L., et al., *Acid multimers of oligodeoxycytidine strands: stoichiometry, base-pair characterization, and proton exchange properties.* Biochemistry, 1993. **32**(23): p. 6019-6031.
43. Rich, A. and D.R. Davies, *A new two stranded helical structure: polyadenylic acid and polyuridylic acid.* Journal of the American Chemical Society, 1956. **78**(14): p. 3548-3549.
44. Rich, A., *A hybrid helix containing both deoxyribose and ribose polynucleotides and its relation to the transfer of information between the nucleic acids.* Proceedings of the National Academy of Sciences, 1960. **46**(8): p. 1044-1053.
45. Herrero-Galán, E., et al., *Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level.* Journal of the American Chemical Society, 2012. **135**(1): p. 122-131.
46. Devi, G., et al., *RNA triplexes: from structural principles to biological and biotech applications.* Wiley Interdisciplinary Reviews: RNA, 2015. **6**(1): p. 111-128.

47. Kumari, S., et al., *An RNA G-quadruplex in the 5′ UTR of the NRAS proto-oncogene modulates translation.* Nature chemical biology, 2007. **3**(4): p. 218-221.
48. Kumari, S., A. Bugaut, and S. Balasubramanian, *Position and stability are determining factors for translation repression by an RNA G-quadruplex-forming sequence within the 5′ UTR of the NRAS proto-oncogene.* Biochemistry, 2008. **47**(48): p. 12664.
49. Bugaut, A. and S. Balasubramanian, *5′-UTR RNA G-quadruplexes: translation regulation and targeting.* Nucleic acids research, 2012. **40**(11): p. 4727-4741.
50. Beaudoin, J.-D. and J.-P. Perreault, *5′-UTR G-quadruplex structures acting as translational repressors.* Nucleic acids research, 2010. **38**(20): p. 7022-7036.
51. Beaudoin, J.-D. and J.-P. Perreault, *Exploring mRNA 3′-UTR G-quadruplexes: evidence of roles in both alternative polyadenylation and mRNA shortening.* Nucleic acids research, 2013. **41**(11): p. 5898-5911.
52. Xu, Y., et al., *Telomeric repeat-containing RNA structure in living cells.* Proceedings of the National Academy of Sciences, 2010. **107**(33): p. 14579-14584.
53. Dykxhoorn, D.M., C.D. Novina, and P.A. Sharp, *Killing the messenger: short RNAs that silence gene expression.* Nature reviews Molecular cell biology, 2003. **4**(6): p. 457-467.
54. Lu, J., et al., *MicroRNA expression profiles classify human cancers.* nature, 2005. **435**(7043): p. 834-838.
55. Bumcrot, D., et al., *RNAi therapeutics: a potential new class of pharmaceutical drugs.* Nature chemical biology, 2006. **2**(12): p. 711-719.
56. Calin, G.A. and C.M. Croce, *MicroRNA signatures in human cancers.* Nature reviews cancer, 2006. **6**(11): p. 857-866.
57. Petersen, C.P., et al., *Short RNAs repress translation after initiation in mammalian cells.* Molecular cell, 2006. **21**(4): p. 533-542.
58. Ventura, A. and T. Jacks, *MicroRNAs and cancer: short RNAs go a long way.* Cell, 2009. **136**(4): p. 586-591.
59. Siomi, M.C., et al., *PIWI-interacting small RNAs: the vanguard of genome defence.* Nature reviews Molecular cell biology, 2011. **12**(4): p. 246-258.
60. Gellert, M., M.N. Lipsett, and D.R. Davies, *Helix formation by guanylic acid.* Proceedings of the National Academy of Sciences, 1962. **48**(12): p. 2013-2018.
61. Lipps, H.J., W. Gruissem, and D.M. Prescott, *Higher order DNA structure in macronuclear chromatin of the hypotrichous ciliate Oxytricha nova.* Proceedings of the National Academy of Sciences, 1982. **79**(8): p. 2495-2499.
62. Sundquist, W.I. and A. Klug, *Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops.* 1989.

63. Parkinson, G.N., M.P. Lee, and S. Neidle, *Crystal structure of parallel quadruplexes from human telomeric DNA.* Nature, 2002. **417**(6891): p. 876-880.

64. Schultze, P., et al., *The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the Oxytricha nova telomere repeat oligonucleotide d (G4T4G4).* Nucleic acids research, 1999. **27**(15): p. 3018-3028.

65. Cevec, M. and J. Plavec, *Role of loop residues and cations on the formation and stability of dimeric DNA G-quadruplexes.* Biochemistry, 2005. **44**(46): p. 15238-15246.

66. Clay, E.H. and I.R. Gould, *A combined QM and MM investigation into guanine quadruplexes.* Journal of Molecular Graphics and Modelling, 2005. **24**(2): p. 138-146.

67. Dingley, A.J., et al., *Characterization of the cation and temperature dependence of DNA quadruplex hydrogen bond properties using high-resolution NMR.* Journal of the American Chemical Society, 2005. **127**(41): p. 14466-14472.

68. van Mourik, T. and A.J. Dingley, *Characterization of the Monovalent Ion Position and Hydrogen-Bond Network in Guanine Quartets by DFT Calculations of NMR Parameters.* Chemistry–A European Journal, 2005. **11**(20): p. 6064-6079.

69. Burge, S., et al., *Quadruplex DNA: sequence, topology and structure.* Nucleic acids research, 2006. **34**(19): p. 5402-5415.

70. Neidle, S., *4 - Nonstandard and Higher-Order DNA Structures: DNA–DNA Recognition*, in *Principles of Nucleic Acid Structure*. 2008, Academic Press: New York. p. 81-131.

71. Aboul-Ela, F., A.I. Murchie, and D.M. Lilley, *NMR study of parallel-stranded tetraplex formation by the hexadeoxynucleotide d (TG4T).* Nature, 1992. **360**(6401): p. 280-282.

72. Wang, Y. and D.J. Patel, *Guanine residues in d (T2AG3) and d (T2G4) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution.* Biochemistry, 1992. **31**(35): p. 8112-8119.

73. Wang, Y. and D.J. Patel, *Solution structure of the human telomeric repeat d [AG3 (T2AG3) 3] G-tetraplex.* Structure, 1993. **1**(4): p. 263-282.

74. Phillips, K., et al., *The crystal structure of a parallel-stranded guanine tetraplex at 0.95 Å resolution.* Journal of molecular biology, 1997. **273**(1): p. 171-182.

75. Horvath, M.P. and S.C. Schultz, *DNA G-quartets in a 1.86 Å resolution structure of an Oxytricha nova telomeric protein-DNA complex.* Journal of molecular biology, 2001. **310**(2): p. 367-377.

76. Haider, S., G.N. Parkinson, and S. Neidle, *Crystal structure of the potassium form of an Oxytricha nova G-quadruplex.* Journal of molecular biology, 2002. **320**(2): p. 189-200.

77. Črnugelj, M., P. Šket, and J. Plavec, *Small change in a G-rich sequence, a dramatic change in topology: new dimeric G-quadruplex folding motif with unique loop orientations*. Journal of the American Chemical Society, 2003. **125**(26): p. 7866-7871.
78. Hazel, P., et al., *Loop-length-dependent folding of G-quadruplexes*. Journal of the American Chemical Society, 2004. **126**(50): p. 16405-16415.
79. Luu, K.N., et al., *Structure of the human telomere in K+ solution: an intramolecular (3+ 1) G-quadruplex scaffold.* Journal of the American Chemical Society, 2006. **128**(30): p. 9963-9970.
80. Rachwal, P.A., et al., *Intramolecular DNA quadruplexes with different arrangements of short and long loops*. Nucleic acids research, 2007. **35**(12): p. 4214-4222.
81. Rachwal, P.A., T. Brown, and K.R. Fox, *Sequence effects of single base loops in intramolecular quadruplex DNA*. FEBS letters, 2007. **581**(8): p. 1657-1660.
82. Phan, A.T., et al., *Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter.* Journal of the American Chemical Society, 2007. **129**(14): p. 4386-4392.
83. Karsisiotis, A.I., et al., *Topological Characterization of Nucleic Acid G-Quadruplexes by UV Absorption and Circular Dichroism.* Angewandte Chemie International Edition, 2011. **50**(45): p. 10645-10648.
84. Cang, X., J. Šponer, and T.E. Cheatham, *Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes*. Nucleic acids research, 2011. **39**(10): p. 4499-4512.
85. Rawal, P., et al., *Genome-wide prediction of G4 DNA as regulatory motifs: role in Escherichia coli global regulation.* Genome research, 2006. **16**(5): p. 644-655.
86. Beaume, N., et al., *Genome-wide study predicts promoter-G4 DNA motifs regulate selective functions in bacteria: radioresistance of D. radiodurans involves G4 DNA-mediated regulation.* Nucleic acids research, 2013. **41**(1): p. 76-89.
87. Johnson, J.E., et al., *In vivo veritas: using yeast to probe the biological functions of G-quadruplexes*. Biochimie, 2008. **90**(8): p. 1250-1263.
88. Huppert, J., *Hunting G-quadruplexes*. Biochimie, 2008. **90**(8): p. 1140-1148.
89. Maizels, N., *Dynamic roles for G4 DNA in the biology of eukaryotic cells*. Nature structural & molecular biology, 2006. **13**(12): p. 1055-1059.
90. Sundquist, W.I. and S. Heaphy, *Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus 1 genomic RNA*. Proceedings of the National Academy of Sciences, 1993. **90**(8): p. 3393-3397.
91. Norseen, J., F.B. Johnson, and P.M. Lieberman, *Role for G-quadruplex RNA binding by Epstein-Barr virus nuclear antigen 1 in DNA replication and metaphase chromosome attachment*. Journal of virology, 2009. **83**(20): p. 10336-10346.

92. Tuesuwan, B., et al., *Simian virus 40 large T-antigen G-quadruplex DNA helicase inhibition by G-quadruplex DNA-interactive agents.* Biochemistry, 2008. **47**(7): p. 1896-1909.
93. Métifiot, M., et al., *G-quadruplexes in viruses: function and potential therapeutic applications.* Nucleic acids research, 2014. **42**(20): p. 12352-12366.
94. Capra, J.A., et al., *G-quadruplex DNA sequences are evolutionarily conserved and associated with distinct genomic features in Saccharomyces cerevisiae.* PLoS Comput Biol, 2010. **6**(7): p. e1000861.
95. Lipps, H.J. and D. Rhodes, *G-quadruplex structures: in vivo evidence and function.* Trends in cell biology, 2009. **19**(8): p. 414-422.
96. Rhodes, D. and H.J. Lipps, *G-quadruplexes and their regulatory roles in biology.* Nucleic acids research, 2015: p. gkv862.
97. de Lange, T., *Protection of mammalian telomeres.* Oncogene, 2002. **21**(4): p. 532.
98. Moyzis, R.K., et al., *A highly conserved repetitive DNA sequence,(TTAGGG) n, present at the telomeres of human chromosomes.* Proceedings of the National Academy of Sciences, 1988. **85**(18): p. 6622-6626.
99. De Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres.* Genes & development, 2005. **19**(18): p. 2100-2110.
100. Blackburn, E.H., *Switching and signaling at the telomere.* Cell, 2001. **106**(6): p. 661-673.
101. Collins, K. and J.R. Mitchell, *Telomerase in the human organism.* Oncogene, 2002. **21**(4): p. 564.
102. Oeseburg, H., et al., *Telomere biology in healthy aging and disease.* Pflügers Archiv-European Journal of Physiology, 2010. **459**(2): p. 259-268.
103. Kim, N.W., et al., *Specific association of human telomerase activity with immortal cells and cancer.* Science, 1994. **266**(5193): p. 2011.
104. Meyerson, M., *Role of telomerase in normal and cancer cells.* Journal of Clinical Oncology, 2000. **18**(13): p. 2626-2634.
105. Hemann, M.T., et al., *The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability.* Cell, 2001. **107**(1): p. 67-77.
106. Blasco, M.A., *Telomeres and human disease: ageing, cancer and beyond.* Nature Reviews Genetics, 2005. **6**(8): p. 611-622.
107. Collado, M., M.A. Blasco, and M. Serrano, *Cellular senescence in cancer and aging.* Cell, 2007. **130**(2): p. 223-233.
108. Shay, J.W. and W.E. Wright. *Role of telomeres and telomerase in cancer.* in *Seminars in cancer biology*. 2011. Elsevier.

109. Hahn, W.C., et al., *Inhibition of telomerase limits the growth of human cancer cells.* Nature medicine, 1999. **5**(10): p. 1164-1170.

110. Griffith, J.D., et al., *Mammalian telomeres end in a large duplex loop.* Cell, 1999. **97**(4): p. 503-514.

111. De Lange, T., *T-loops and the origin of telomeres.* Nature Reviews Molecular Cell Biology, 2004. **5**(4): p. 323-329.

112. Paeschke, K., et al., *Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo.* Nature structural & molecular biology, 2005. **12**(10): p. 847-854.

113. Tang, J., et al., *G-quadruplex preferentially forms at the very 3′ end of vertebrate telomeric DNA.* Nucleic acids research, 2008. **36**(4): p. 1200-1208.

114. Biffi, G., et al., *Quantitative visualization of DNA G-quadruplex structures in human cells.* Nature chemistry, 2013. **5**(3): p. 182-186.

115. Yang, Q., et al., *Verification of specific G-quadruplex structure by using a novel cyanine dye supramolecular assembly: I. Recognizing mixed G-quadruplex in human telomeres.* Chemical Communications, 2009(9): p. 1103-1105.

116. Azzalin, C.M., et al., *Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends.* Science, 2007. **318**(5851): p. 798-801.

117. Schoeftner, S. and M.A. Blasco, *Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II.* Nature cell biology, 2008. **10**(2): p. 228-236.

118. Porro, A., et al., *Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways.* Molecular and cellular biology, 2010. **30**(20): p. 4808-4817.

119. Deng, Z., et al., *TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres.* Molecular cell, 2009. **35**(4): p. 403-413.

120. De Silanes, I.L., M.S. d'Alcontres, and M.A. Blasco, *TERRA transcripts are bound by a complex array of RNA-binding proteins.* Nature communications, 2010. **1**: p. 33.

121. Sun, D., et al., *Inhibition of human telomerase by a G-quadruplex-interactive compound.* Journal of medicinal chemistry, 1997. **40**(14): p. 2113-2116.

122. Paeschke, K., K.R. McDonald, and V.A. Zakian, *Telomeres: structures in need of unwinding.* FEBS letters, 2010. **584**(17): p. 3760-3772.

123. Sharma, S., K.M. Doherty, and R.M. Brosh, *Mechanisms of RecQ helicases in pathways of DNA metabolism and maintenance of genomic stability.* Biochemical Journal, 2006. **398**(3): p. 319-337.

124. Paeschke, K., et al., *Telomerase recruitment by the telomere end binding protein-β facilitates G-quadruplex DNA unfolding in ciliates.* Nature structural & molecular biology, 2008. **15**(6): p. 598-604.
125. Chavez, A., A.M. Tsou, and F.B. Johnson, *Telomeres do the (un) twist: helicase actions at chromosome termini.* Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 2009. **1792**(4): p. 329-340.
126. Postberg, J., et al., *A telomerase-associated RecQ protein-like helicase resolves telomeric G-quadruplex structures during replication.* Gene, 2012. **497**(2): p. 147-154.
127. A Juranek, S. and K. Paeschke, *Cell cycle regulation of G-quadruplex DNA structures at telomeres.* Current pharmaceutical design, 2012. **18**(14): p. 1867-1872.
128. Brosh Jr, R.M., *DNA helicases involved in DNA repair and their roles in cancer.* Nature Reviews Cancer, 2013. **13**(8): p. 542-558.
129. Wickramasinghe, C.M., et al., *Contributions of the specialised DNA polymerases to replication of structured DNA.* DNA repair, 2015. **29**: p. 83-90.
130. Haeusler, A.R., et al., *C9orf72 nucleotide repeat structures initiate molecular cascades of disease.* Nature, 2014. **507**(7491): p. 195-200.
131. Mergny, J.-L. and C. Hélène, *G-quadruplex DNA: a target for drug design.* Nature medicine, 1998. **4**(12): p. 1366-1367.
132. Ou, T.m., et al., *G-Quadruplexes: Targets in Anticancer Drug Design.* ChemMedChem, 2008. **3**(5): p. 690-713.
133. Wheelhouse, R.T., et al., *Cationic porphyrins as telomerase inhibitors: the interaction of tetra-(N-methyl-4-pyridyl) porphine with quadruplex DNA.* Journal of the American Chemical Society, 1998. **120**(13): p. 3261-3262.
134. Haider, S.M., G.N. Parkinson, and S. Neidle, *Structure of a G-quadruplex–ligand complex.* Journal of molecular biology, 2003. **326**(1): p. 117-125.
135. Gavathiotis, E., et al., *Drug recognition and stabilisation of the parallel-stranded DNA quadruplex d (TTAGGGT) 4 containing the human telomeric repeat.* Journal of molecular biology, 2003. **334**(1): p. 25-36.
136. Schouten, J.A., et al., *G-quadruplex-specific peptide− hemicyanine ligands by partial combinatorial selection.* Journal of the American Chemical Society, 2003. **125**(19): p. 5594-5595.
137. Kerwin, S.M., *G-quadruplex DNA as a target for drug design.* Current pharmaceutical design, 2000. **6**(4): p. 441-471.
138. Cuesta, J., M.A. Read, and S. Neidle, *The design of G-quadruplex ligands as telomerase inhibitors.* Mini reviews in medicinal chemistry, 2003. **3**(1): p. 11-21.

139. Guittat, L., et al., *Targeting human telomerase for cancer therapeutics.* Cytotechnology, 2004. **45**(1-2): p. 75-90.

140. Pendino, F., et al., *Telomeres and telomerase: Pharmacological targets for new anticancer strategies?* Current cancer drug targets, 2006. **6**(2): p. 147-180.

141. Riou, J., et al., *Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands.* Proceedings of the National Academy of Sciences, 2002. **99**(5): p. 2672-2677.

142. Tauchi, T., et al., *Activity of a novel G-quadruplex-interactive telomerase inhibitor, telomestatin (SOT-095), against human leukemia cells: involvement of ATM-dependent DNA damage response pathways.* Oncogene, 2003. **22**(34): p. 5338-5347.

143. Pennarun, G., et al., *Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands.* Oncogene, 2005. **24**(18): p. 2917-2928.

144. Zhou, J., et al., *Senescence and telomere shortening induced by novel potent G-quadruplex interactive agents, quindoline derivatives, in human cancer cell lines.* Oncogene, 2006. **25**(4): p. 503-511.

145. Gomez, D., et al., *Interaction of telomestatin with the telomeric single-strand overhang.* Journal of Biological Chemistry, 2004. **279**(40): p. 41487-41494.

146. Neidle, S. and G. Parkinson, *Telomere maintenance as a target for anticancer drug discovery.* Nature Reviews Drug Discovery, 2002. **1**(5): p. 383-393.

147. Marsh, T.C. and E. Henderson, *G-wires: self-assembly of a telomeric oligonucleotide, d (GGGGTTGGGG), into large superstructures.* Biochemistry, 1994. **33**(35): p. 10718-10724.

148. Marsh, T.C., J. Vesenka, and E. Henderson, *A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy.* Nucleic acids research, 1995. **23**(4): p. 696-700.

149. Calzolari, A., et al., *G-quartet biomolecular nanowires.* Applied physics letters, 2002. **80**(18): p. 3331-3333.

150. Sondermann, A., et al. *Assembly Of G-Quartet Based DNA Superstructures (G-Wires).* in *AIP Conference Proceedings*. 2002. AIP.

151. Davis, J.T., *G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry.* Angewandte Chemie International Edition, 2004. **43**(6): p. 668-698.

152. Kotlyar, A.B., et al., *Long, Monomolecular Guanine-Based Nanowires.* Advanced Materials, 2005. **17**(15): p. 1901-1905.

153. Davis, J.T. and G.P. Spada, *Supramolecular architectures generated by self-assembly of guanosine derivatives.* Chemical Society Reviews, 2007. **36**(2): p. 296-313.

154. Liu, S.P., et al., *Direct Measurement of Electrical Transport Through G-Quadruplex DNA with Mechanically Controllable Break Junction Electrodes.* Angewandte Chemie International Edition, 2010. **49**(19): p. 3313-3316.

155. Kotia, R.B., L. Li, and L.B. McGown, *Separation of nontarget compounds by DNA aptamers.* Analytical Chemistry, 2000. **72**(4): p. 827-831.

156. Rehder, M.A. and L.B. McGown, *Open-tubular capillary electrochromatography of bovine β-lactoglobulin variants A and B using an aptamer stationary phase.* Electrophoresis, 2001. **22**(17): p. 3759-3764.

157. Charles, J.A. and L.B. McGown, *Separation of Trp-Arg and Arg-Trp using G-quartet-forming DNA oligonucleotides in open-tubular capillary electrochromatography.* Electrophoresis, 2002. **23**(11): p. 1599-1604.

158. Roe, S., et al., *Tuneable DNA-based asymmetric catalysis using a G-quadruplex supramolecular assembly.* Chemical communications, 2010. **46**(24): p. 4309-4311.

159. Wang, C., et al., *Enantioselective Diels–Alder Reactions with G-Quadruplex DNA-Based Catalysts.* Angewandte Chemie International Edition, 2012. **51**(37): p. 9352-9355.

160. German, I., D.D. Buchanan, and R.T. Kennedy, *Aptamers as ligands in affinity probe capillary electrophoresis.* Analytical Chemistry, 1998. **70**(21): p. 4540-4545.

161. Potyrailo, R.A., et al., *Adapting selected nucleic acid ligands (aptamers) to biosensors.* Analytical Chemistry, 1998. **70**(16): p. 3419-3425.

162. Hamaguchi, N., A. Ellington, and M. Stanton, *Aptamer beacons for the direct detection of proteins.* Analytical biochemistry, 2001. **294**(2): p. 126-131.

163. Li, J.J., X. Fang, and W. Tan, *Molecular aptamer beacons for real-time protein recognition.* Biochemical and biophysical research communications, 2002. **292**(1): p. 31-40.

164. Ueyama, H., M. Takagi, and S. Takenaka, *A novel potassium sensing in aqueous media with a synthetic oligonucleotide derivative. Fluorescence resonance energy transfer associated with guanine quartet− potassium ion complex formation.* Journal of the American Chemical Society, 2002. **124**(48): p. 14286-14287.

165. Song, S., et al., *Aptamer-based biosensors.* TrAC Trends in Analytical Chemistry, 2008. **27**(2): p. 108-117.

166. Guo, L., et al., *A G-quadruplex based label-free fluorescent biosensor for lead ion.* Biosensors and Bioelectronics, 2012. **35**(1): p. 123-127.

167. Yue, Q., et al., *Construction of a controllable förster resonance energy transfer system based on G-quadruplex for DNA sensing.* Biosensors and Bioelectronics, 2013. **40**(1): p. 75-81.

168. Bustamante, C., J.C. Macosko, and G.J. Wuite, *Grabbing the cat by the tail: manipulating molecules one by one.* Nature Reviews Molecular Cell Biology, 2000. **1**(2): p. 130-136.

169. Rief, M., et al., *Reversible unfolding of individual titin immunoglobulin domains by AFM.* science, 1997. **276**(5315): p. 1109-1112.

170. Carrion-Vazquez, M., et al., *Atomic force microscopy captures length phenotypes in single proteins.* Proceedings of the National Academy of Sciences, 1999. **96**(20): p. 11288-11292.

171. Carrion-Vazquez, M., et al., *Mechanical and chemical unfolding of a single protein: a comparison.* Proceedings of the National Academy of Sciences, 1999. **96**(7): p. 3694-3699.

172. Marszalek, P.E., et al., *Mechanical unfolding intermediates in titin modules.* Nature, 1999. **402**(6757): p. 100-103.

173. Socci, N.D., J.N. Onuchic, and P.G. Wolynes, *Stretching lattice models of protein folding.* Proceedings of the National Academy of Sciences, 1999. **96**(5): p. 2031-2035.

174. Yang, G., et al., *Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme.* Proceedings of the National Academy of Sciences, 2000. **97**(1): p. 139-144.

175. Fernandez, J.M. and H. Li, *Force-clamp spectroscopy monitors the folding trajectory of a single protein.* Science, 2004. **303**(5664): p. 1674-1678.

176. Chung, H.S., et al., *Single-molecule fluorescence experiments determine protein folding transition path times.* Science, 2012. **335**(6071): p. 981-984.

177. Erie, D.A., et al., *DNA bending by Cro protein in specific and nonspecific complexes: implications for protein site recognition and specificity.* Science, 1994. **266**(5190): p. 1562.

178. Wang, H., et al., *DNA bending and unbending by MutS govern mismatch recognition and specificity.* Proceedings of the National Academy of Sciences, 2003. **100**(25): p. 14822-14827.

179. Radmacher, M., et al., *Direct observation of enzyme activity with the atomic force microscope.* SCIENCE-NEW YORK THEN WASHINGTON-, 1994: p. 1577-1577.

180. Wuite, G.J., et al., *Single-molecule studies of the effect of template tension on T7 DNA polymerase activity.* Nature, 2000. **404**(6773): p. 103-106.

181. Kishino, A. and T. Yanagida, *Force measurements by micromanipulation of a single actin filament by glass needles.* 1988.

182. Svoboda, K., et al., *Direct observation of kinesin stepping by optical trapping interferometry.* Nature, 1993. **365**(6448): p. 721.

183. Junge, W., *ATP synthase and other motor proteins.* Proceedings of the National Academy of Sciences, 1999. **96**(9): p. 4735-4737.
184. Abbondanzieri, E.A., et al., *Direct observation of base-pair stepping by RNA polymerase.* Nature, 2005. **438**(7067): p. 460-465.
185. Morin, J.A., et al., *Active DNA unwinding dynamics during processive DNA replication.* Proceedings of the National Academy of Sciences, 2012. **109**(21): p. 8115-8120.
186. Morin, J.A., et al., *Mechano-chemical kinetics of DNA replication: identification of the translocation step of a replicative DNA polymerase.* Nucleic acids research, 2015. **43**(7): p. 3643-3652.
187. Morin, J.A., et al., *DNA synthesis determines the binding mode of the human mitochondrial single-stranded DNA-binding protein.* Nucleic Acids Research, 2017.
188. Smith, S.B., L. Finzi, and C. Bustamante, *Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads.* Science, 1992. **258**(5085): p. 1122-1126.
189. Bustamante, C., et al., *Entropic elasticity of lambda-phage DNA.* SCIENCE-NEW YORK THEN WASHINGTON-, 1994: p. 1599-1599.
190. Marko, J.F. and E.D. Siggia, *Stretching dna.* Macromolecules, 1995. **28**(26): p. 8759-8770.
191. Wang, M.D., et al., *Stretching DNA with optical tweezers.* Biophysical journal, 1997. **72**(3): p. 1335-1346.
192. Bustamante, C., et al., *Single-molecule studies of DNA mechanics.* Current opinion in structural biology, 2000. **10**(3): p. 279-285.
193. Mehta, A.D., et al., *Single-molecule biomechanics with optical methods.* Science, 1999. **283**(5408): p. 1689-1695.
194. Clausen-Schaumann, H., et al., *Force spectroscopy with single bio-molecules.* Current opinion in chemical biology, 2000. **4**(5): p. 524-530.
195. Ishii, Y. and T. Yanagida, *Single molecule detection in life sciences.* Single Molecules, 2000. **1**(1): p. 5-16.
196. Bockelmann, U., *Single-molecule manipulation of nucleic acids.* Current opinion in structural biology, 2004. **14**(3): p. 368-373.
197. Neuman, K.C. and A. Nagy, *Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy.* Nature methods, 2008. **5**(6): p. 491.
198. Neuman, K.C., et al., *Characterization of photodamage to Escherichia coli in optical traps.* Biophysical journal, 1999. **77**(5): p. 2856-2863.

199. Bennink, M.L., et al., *Unfolding individual nucleosomes by stretching single chromatin fibers with optical tweezers.* Nature Structural & Molecular Biology, 2001. **8**(7): p. 606-610.

200. Cherney, D.P., T.E. Bridges, and J.M. Harris, *Optical trapping of unilamellar phospholipid vesicles: investigation of the effect of optical forces on the lipid membrane shape by confocal-Raman microscopy.* Analytical chemistry, 2004. **76**(17): p. 4920-4928.

201. Sacconi, L., et al., *Optical micromanipulations inside yeast cells.* Applied optics, 2005. **44**(11): p. 2001-2007.

202. Ashkin, A., et al., *Observation of a single-beam gradient force optical trap for dielectric particles.* Optics letters, 1986. **11**(5): p. 288-290.

203. Arias-González, J. and M. Nieto-Vesperinas, *Optical forces on small particles: attractive and repulsive nature and plasmon-resonance conditions.* JOSA A, 2003. **20**(7): p. 1201-1209.

204. Ashkin, A. and J.M. Dziedzic, *Optical trapping and manipulation of viruses and bacteria.* Science, 1987. **235**: p. 1517-1521.

205. Ashkin, A., J.M. Dziedzic, and T. Yamane, *Optical trapping and manipulation of single cells using infrared laser beams.* Nature, 1987. **330**(6150): p. 769-771.

206. Ashkin, A. and J. Dziedzic, *Internal cell manipulation using infrared laser traps.* Proceedings of the National Academy of Sciences, 1989. **86**(20): p. 7914-7918.

207. Hormeño, S., et al., *Single centrosome manipulation reveals its electric charge and associated dynamic structure.* Biophysical journal, 2009. **97**(4): p. 1022-1030.

208. Arias-Gonzalez, J.R., *Optical tweezers to study viruses*, in *Structure and Physics of Viruses*. 2013, Springer. p. 273-304.

209. Hormeno, S. and J.R. Arias-Gonzalez, *Exploring mechanochemical processes in the cell with optical tweezers.* Biology of the Cell, 2006. **98**(12): p. 679-695.

210. Smith, S.B., Y. Cui, and C. Bustamante, *Optical-trap force transducer that operates by direct measurement of light momentum.* Methods in enzymology, 2003. **361**: p. 134-162.

211. Fällman, E. and O. Axner, *Influence of a glass-water interface on the on-axis trapping of micrometer-sized spherical objects by optical tweezers.* Applied optics, 2003. **42**(19): p. 3915-3926.

212. Neuman, K.C., E.A. Abbondanzieri, and S.M. Block, *Measurement of the effective focal shift in an optical trap.* Optics letters, 2005. **30**(11): p. 1318-1320.

213. Bustamante, C.J. and S.B. Smith, *Light-force sensor and method for measuring axial optical-trap forces from changes in light momentum along an optic axis*. 2006, U.S. Patent No. 7133132 B2

214. de Lorenzo, S., et al., *A temperature-jump optical trap for single-molecule manipulation.* Biophysical journal, 2015. **108**(12): p. 2854-2864.

215. Selvam, S., et al., *Quantification of topological coupling between DNA superhelicity and G-quadruplex formation.* Journal of the American Chemical Society, 2014. **136**(40): p. 13967-13970.

216. Capitanio, M., et al., *FIONA in the trap: the advantages of combining optical tweezers and fluorescence.* Journal of Optics A: Pure and Applied Optics, 2007. **9**(8): p. S157.

217. Cluzel, P., et al., *DNA: an extensible molecule.* Science, 1996. **271**(5250): p. 792.

218. Smith, S.B., Y. Cui, and C. Bustamente, *Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules.* Science, 1996. **271**(5250): p. 795.

219. Dhakal, S., et al., *G-quadruplex and i-motif are mutually exclusive in ILPR double-stranded DNA.* Biophysical journal, 2012. **102**(11): p. 2575-2584.

220. Dhakal, S., et al., *Structural and mechanical properties of individual human telomeric G-quadruplexes in molecularly crowded solutions.* Nucleic acids research, 2013. **41**(6): p. 3915-3923.

221. Garavís, M., et al., *Mechanical unfolding of long human telomeric RNA (TERRA).* Chemical Communications, 2013. **49**(57): p. 6397-6399.

222. Yangyuoru, P.M., et al., *Mechanochemical Properties of Individual Human Telomeric RNA (TERRA) G-Quadruplexes.* ChemBioChem, 2013. **14**(15): p. 1931-1935.

223. Abraham Punnoose, J., et al., *Interaction of G-quadruplexes in the full-length 3' human telomeric overhang.* Journal of the American Chemical Society, 2014. **136**(52): p. 18062-18069.

224. Yu, Z. and H. Mao, *Non-B DNA structures show diverse conformations and complex transition kinetics comparable to RNA or proteins--a perspective from mechanical unfolding and refolding experiments.* Chem Rec, 2013. **13**(1): p. 102-16.

225. Williams, M.C., et al., *Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting.* Biophysical Journal, 2001. **80**(2): p. 874-881.

226. Wenner, J.R., et al., *Salt dependence of the elasticity and overstretching transition of single DNA molecules.* Biophysical journal, 2002. **82**(6): p. 3160-3169.

227. van Mameren, J., et al., *Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging.* Proceedings of the National Academy of Sciences, 2009. **106**(43): p. 18231-18236.

228. Williams, M.C., I. Rouzina, and M.J. McCauley, *Peeling back the mystery of DNA overstretching.* Proceedings of the National Academy of Sciences, 2009. **106**(43): p. 18047-18048.

229. Fu, H., et al., *Two distinct overstretched DNA states.* Nucleic acids research, 2010: p. gkq309.

230. Fu, H., et al., *Transition dynamics and selection of the distinct S-DNA and strand unpeeling modes of double helix overstretching.* Nucleic acids research, 2010: p. gkq1278.

231. Paik, D.H. and T.T. Perkins, *Overstretching DNA at 65 pN does not require peeling from free ends or nicks.* Journal of the American Chemical Society, 2011. **133**(10): p. 3219-3221.

232. Bosaeus, N., et al., *Tension induces a base-paired overstretched DNA conformation.* Proceedings of the National Academy of Sciences, 2012. **109**(38): p. 15179-15184.

233. Zhang, X., et al., *Two distinct overstretched DNA structures revealed by single-molecule thermodynamics measurements.* Proceedings of the National Academy of Sciences, 2012. **109**(21): p. 8103-8108.

234. King, G.A., et al., *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching using fluorescence microscopy.* Proceedings of the National Academy of Sciences, 2013. **110**(10): p. 3859-3864.

235. Zhang, X., et al., *Revealing the competition between peeled ssDNA, melting bubbles, and S-DNA during DNA overstretching by single-molecule calorimetry.* Proceedings of the National Academy of Sciences, 2013. **110**(10): p. 3865-3870.

236. Peterman, E.J., F. Gittes, and C.F. Schmidt, *Laser-induced heating in optical traps.* Biophysical journal, 2003. **84**(2): p. 1308-1316.

237. Seol, Y., A.E. Carpenter, and T.T. Perkins, *Gold nanoparticles: enhanced optical trapping and sensitivity coupled with significant heating.* Optics letters, 2006. **31**(16): p. 2429-2431.

238. Hormeño, S., et al., *Laser Heating Tunability by Off-Resonant Irradiation of Gold Nanoparticles.* Small, 2014. **10**(2): p. 376-384.

239. Hormeno, S. and J.R. Arias-Gonzalez, *Exploring mechanochemical processes in the cell with optical tweezers.* Biol Cell, 2006. **98**(12): p. 679-95.

240. Binnig, G., N. Garcia, and H. Rohrer, *Conductivity sensitivity of inelastic scanning tunneling microscopy.* Physical Review B, 1985. **32**(2): p. 1336.

241. Binnig, G., C.F. Quate, and C. Gerber, *Atomic force microscope.* Physical review letters, 1986. **56**(9): p. 930.

242. Shao, Z., J. Yang, and A.P. Somlyo, *Biological atomic force microscopy: from microns to nanometers and beyond.* Annual review of cell and developmental biology, 1995. **11**(1): p. 241-265.

243. Bustamante, C., C. Rivetti, and D.J. Keller, *Scanning force microscopy under aqueous solutions.* Current opinion in structural biology, 1997. **7**(5): p. 709-716.

244. Hugel, T. and M. Seitz, *The study of molecular interactions by AFM force spectroscopy.* Macromolecular rapid communications, 2001. **22**(13): p. 989-1016.

245. Eaton, P. and P. West, *Atomic force microscopy*. 2010: Oxford University Press.

246. Faraldos, M. and C. Goberna, *Técnicas de análisis y caracterización de materiales*. 2002: Consejo Superior de Investigaciones Científicas.

247. Allen, M.J., et al., *Atomic force microscope measurements of nucleosome cores assembled along defined DNA sequences.* Biochemistry, 1993. **32**(33): p. 8390-8396.

248. Bustamante, C., D. Keller, and G. Yang, *Scanning force microscopy of nucleic acids and nucleoprotein assemblies*. Current Opinion in Structural Biology, 1993. **3**(3): p. 363-372.

249. Hansma, H.G., et al., *Atomic force microscopy of DNA in aqueous solutions*. Nucleic acids research, 1993. **21**(3): p. 505-512.

250. Rees, W.A., et al., *Evidence of DNA bending in transcription complexes imaged by scanning force microscopy.* Science, 1993. **260**(5114): p. 1646-1650.

251. Leuba, S.H., et al., *Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy.* Proceedings of the National Academy of Sciences, 1994. **91**(24): p. 11621-11625.

252. Bustamante, C., D.A. Erie, and D. Keller, *Biochemical and structural applications of scanning force microscopy.* Current Opinion in Structural Biology, 1994. **4**(5): p. 750-760.

253. Hansma, H., et al., *Properties of biomolecules measured from atomic force microscope images: a review.* Journal of structural biology, 1997. **119**(2): p. 99-108.

254. Costa, L., et al., *Structural studies of oligonucleotides containing G-quadruplex motifs using AFM.* Biochemical and biophysical research communications, 2004. **313**(4): p. 1065-1072.

255. Moreno-Herrero, F., et al., *Atomic force microscopy contact, tapping, and jumping modes for imaging biological samples in liquids*. Physical Review E, 2004. **69**(3): p. 031915.

256. Xu, Y. and M. Komiyama. *The structural studies of human telomeric DNA using AFM*. in *Nucleic acids symposium series*. 2007. Oxford Univ Press.

257. Neaves, K.J., et al., *Direct visualization of G-quadruplexes in DNA using atomic force microscopy.* Nucleic acids research, 2009. **37**(18): p. 6269-6275.

258. Xu, Y., et al., *Consecutive Formation of G-Quadruplexes in Human Telomeric-Overhang DNA: A Protective Capping Structure for Telomere Ends*. Angewandte Chemie International Edition, 2009. **48**(42): p. 7833-7836.

259. Mela, I., et al., *Demonstration of ligand decoration, and ligand-induced perturbation, of G-quadruplexes in a plasmid using atomic force microscopy.* Biochemistry, 2012. **51**(2): p. 578-585.

260. Fuentes-Perez, M.E., M.S. Dillingham, and F. Moreno-Herrero, *AFM volumetric methods for the characterization of proteins and nucleic acids.* Methods, 2013. **60**(2): p. 113-121.

261. Ares, P., et al., *High resolution atomic force microscopy of double-stranded RNA*. Nanoscale, 2016. **8**(23): p. 11818-11826.

262. Luke, B., et al., *The Rat1p 5′ to 3′ exonuclease degrades telomeric repeat-containing RNA and promotes telomere elongation in Saccharomyces cerevisiae.* Molecular cell, 2008. **32**(4): p. 465-477.

263. Caslini, C., et al., *MLL associates with telomeres and regulates telomeric repeat-containing RNA transcription.* Molecular and cellular biology, 2009. **29**(16): p. 4519-4526.

264. Redon, S., P. Reichenbach, and J. Lingner, *The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase.* Nucleic acids research, 2010: p. gkq296.

265. Deng, Z., A.E. Campbell, and P.M. Lieberman, *TERRA, CpG methylation, and telomere heterochromatin: lessons from ICF syndrome cells*. Cell Cycle, 2010. **9**(1): p. 69-74.

266. Horard, B. and E. Gilson, *Telomeric RNA enters the game*. Nature cell biology, 2008. **10**(2): p. 113-114.

267. Luke, B. and J. Lingner, *TERRA: telomeric repeat-containing RNA*. The EMBO journal, 2009. **28**(17): p. 2503-2510.

268. Feuerhahn, S., et al., *TERRA biogenesis, turnover and implications for function.* FEBS letters, 2010. **584**(17): p. 3812-3818.

269. Xu, Y., K. Kaminaga, and M. Komiyama, *G-quadruplex formation by human telomeric repeats-containing RNA in Na+ solution.* Journal of the American Chemical Society, 2008. **130**(33): p. 11179-11184.

270. Randall, A. and J.D. Griffith, *Structure of long telomeric RNA transcripts the G-rich rna forms a compact repeating structure containing G-quartets*. Journal of Biological Chemistry, 2009. **284**(21): p. 13980-13986.
271. Phan, A.T., *Human telomeric G-quadruplex: structures of DNA and RNA sequences.* The FEBS journal, 2010. **277**(5): p. 1107-1117.
272. Mullis, K.B., et al., *Process for amplifying, detecting, and/or-cloning nucleic acid sequences*. 1987, Google Patents.
273. Horcas, I., et al., *WSXM: a software for scanning probe microscopy and a tool for nanotechnology.* Review of Scientific Instruments, 2007. **78**(1): p. 013705.
274. Egli, M. and W. Saenger, *Principles of nucleic acid structure*. 2013: Springer Science & Business Media.
275. Laughlan, G., et al., *The high-resolution crystal structure of a parallel-stranded guanine tetraplex.* Science, 1994. **265**(5171): p. 520-524.
276. Moreno-Herrero, F., J. Colchero, and A. Baro, *DNA height in scanning force microscopy.* Ultramicroscopy, 2003. **96**(2): p. 167-174.
277. Bustamante, C., et al., *Mechanical processes in biochemistry.* Annual review of biochemistry, 2004. **73**(1): p. 705-748.
278. Tinoco, I., G. Chen, and X. Qu, *RNA reactions one molecule at a time.* Cold Spring Harbor perspectives in biology, 2010. **2**(11): p. a003624.
279. Gray, R.D., J.O. Trent, and J.B. Chaires, *Folding and unfolding pathways of the human telomeric G-quadruplex.* J Mol Biol, 2014. **426**(8): p. 1629-50.
280. Mashimo, T., et al., *Folding pathways of human telomeric type-1 and type-2 G-quadruplex structures.* J Am Chem Soc, 2010. **132**(42): p. 14910-8.
281. Rajendran, A., et al., *Direct and single-molecule visualization of the solution-state structures of G-hairpin and G-triplex intermediates.* Angew Chem Int Ed Engl, 2014. **53**(16): p. 4107-12.
282. Bustamante, C., *Unfolding single RNA molecules: bridging the gap between equilibrium and non-equilibrium statistical thermodynamics.* Quarterly Reviews of Biophysics, 2005. **38**(04): p. 291-301.
283. Abraham Punnoose, J., et al., *Interaction of G-quadruplexes in the full-length 3' human telomeric overhang.* J Am Chem Soc, 2014. **136**(52): p. 18062-9.
284. Yangyuoru, P.M., et al., *Mechanochemical properties of individual human telomeric RNA (TERRA) G-quadruplexes.* Chembiochem, 2013. **14**(15): p. 1931-5.
285. de Messieres, M., et al., *Single-molecule study of G-quadruplex disruption using dynamic force spectroscopy.* Phys Rev Lett, 2012. **109**(5): p. 058101.

286. Yu, Z., et al., *ILPR G-quadruplexes formed in seconds demonstrate high mechanical stabilities.* J Am Chem Soc, 2009. **131**(5): p. 1876-82.

287. Bergues-Pupo, A.E., et al., *Role of the central cations in the mechanical unfolding of DNA and RNA G-quadruplexes.* Nucleic Acids Res, 2015. **43**(15): p. 7638-47.

288. Lorenz, R., et al., *2D meets 4G: G-quadruplexes in RNA secondary structure prediction.* IEEE/ACM Trans Comput Biol Bioinform, 2013. **10**(4): p. 832-44.

289. Bugaut, A., P. Murat, and S. Balasubramanian, *An RNA hairpin to G-quadruplex conformational transition.* J Am Chem Soc, 2012. **134**(49): p. 19953-6.

290. Hardin, C.C., et al., *Cation-dependent transition between the quadruplex and Watson-Crick hairpin forms of d(CGCG3GCG).* Biochemistry, 1992. **31**(3): p. 833-41.

291. Kuo, M.H., et al., *Conformational transition of a hairpin structure to G-quadruplex within the WNT1 gene promoter.* J Am Chem Soc, 2015. **137**(1): p. 210-8.

292. de Lorenzo, S., et al., *A Temperature-Jump Optical Trap for Single-Molecule Manipulation.* Biophys J, 2015. **108**(12): p. 2854-64.

293. http://tweezerslab.unipr.it/cgi-bin/mt/home.pl.

294. Garavis, M., et al., *Mechanical unfolding of long human telomeric RNA (TERRA).* Chem Commun (Camb), 2013. **49**(57): p. 6397-9.

295. Bustamante, C., *In singulo biochemistry: when less is more.* Annu Rev Biochem, 2008. **77**: p. 45-50.

296. Arias-Gonzalez, J.R., *Single-molecule portrait of DNA and RNA double helices.* Integr Biol (Camb), 2014. **6**(10): p. 904-25.

297. Herrero-Galan, E., et al., *Mechanical identities of RNA and DNA double helices unveiled at the single-molecule level.* J Am Chem Soc, 2013. **135**(1): p. 122-31.

298. Martadinata, H. and A.T. Phan, *Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K+ solution.* J Am Chem Soc, 2009. **131**(7): p. 2570-8.

299. Alemany, A. and F. Ritort, *Determination of the elastic properties of short ssDNA molecules by mechanically folding and unfolding DNA hairpins.* Biopolymers, 2014. **101**(12): p. 1193-9.

300. Saenger, W., *Principles of nucleic acid structure*. 2nd ed. 1984: Springer-Verlag.

301. Bizarro, C.V., A. Alemany, and F. Ritort, *Non-specific binding of Na+ and Mg2+ to RNA determined by force spectroscopy methods.* Nucleic Acids Res, 2012. **40**(14): p. 6922-35.

302. Bosco, A., J. Camunas-Soler, and F. Ritort, *Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions.* Nucleic Acids Res, 2014. **42**(3): p. 2064-74.

303. Dessinges, M.-N., et al., *Stretching single stranded DNA, a model polyelectrolyte.* Physical review letters, 2002. **89**(24): p. 248102.

304. Zhang, Z., et al., *Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: insights into the interconversion of human telomeric G-quadruplex structures.* Nucleic Acids Res, 2010. **38**(3): p. 1009-21.

305. Li, W., et al., *Direct measurement of sequential folding pathway and energy landscape of human telomeric G-quadruplex structures.* J Am Chem Soc, 2013. **135**(17): p. 6423-6.

306. Zhang, A.Y. and S. Balasubramanian, *The kinetics and folding pathways of intramolecular G-quadruplex nucleic acids.* J Am Chem Soc, 2012. **134**(46): p. 19297-308.

307. You, H., et al., *Dynamics and stability of polymorphic human telomeric G-quadruplex under tension.* Nucleic Acids Res, 2014. **42**(13): p. 8789-95.

308. Long, X., et al., *Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy.* Nucleic Acids Res, 2013. **41**(4): p. 2746-55.

309. Koirala, D., et al., *Long-loop G-quadruplexes are misfolded population minorities with fast transition kinetics in human telomeric sequences.* J Am Chem Soc, 2013. **135**(6): p. 2235-41.

310. Koirala, D., et al., *Intramolecular folding in three tandem guanine repeats of human telomeric DNA.* Chem Commun (Camb), 2012. **48**(14): p. 2006-8.

311. Hsin, J., et al., *Molecular origin of the hierarchical elasticity of titin: simulation, experiment, and theory.* Annu Rev Biophys, 2011. **40**: p. 187-203.

312. Jacobson, D.R., D.B. McIntosh, and O.A. Saleh, *The snakelike chain character of unstructured RNA.* Biophys J, 2013. **105**(11): p. 2569-76.

313. Dessinges, M.N., et al., *Stretching single stranded DNA, a model polyelectrolyte.* Phys Rev Lett, 2002. **89**(24): p. 248102.

314. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Res, 2003. **31**(13): p. 3406-15.

315. Collie, G.W. and G.N. Parkinson, *The application of DNA and RNA G-quadruplexes to therapeutic medicines.* Chemical Society Reviews, 2011. **40**(12): p. 5867-5892.

316. Boncina, M., et al., *Energetic basis of human telomeric DNA folding into G-quadruplex structures.* J Am Chem Soc, 2012. **134**(23): p. 9657-63.

317. Jarzynski, C., *Nonequilibrium Equality for Free Energy Differences.* Physical Review Letters, 1997. **78**(14): p. 2690-2693.

318. Mergny, J.L., A.T. Phan, and L. Lacroix, *Following G-quartet formation by UV-spectroscopy*. FEBS Lett, 1998. **435**(1): p. 74-8.
319. Yurenko, Y.P., et al., *Exploring non-covalent interactions in guanine-and xanthine-based model DNA quadruplex structures: a comprehensive quantum chemical approach.* Physical Chemistry Chemical Physics, 2014. **16**(5): p. 2072-2084.
320. Poudel, L., et al., *Implication of the solvent effect, metal ions and topology in the electronic structure and hydrogen bonding of human telomeric G-quadruplex DNA*. Physical Chemistry Chemical Physics, 2016. **18**(31): p. 21573-21585.
321. Li, H., E.-h. Cao, and T. Gisler, *Force-induced unfolding of human telomeric G-quadruplex: a steered molecular dynamics simulation study.* Biochemical and biophysical research communications, 2009. **379**(1): p. 70-75.
322. Li, M.-H., et al., *Toward a full structural characterization of G-quadruplex DNA in aqueous solution: Molecular dynamics simulations of four G-quadruplex molecules.* Journal of Molecular Structure: THEOCHEM, 2010. **952**(1): p. 96-102.
323. Yang, C., S. Jang, and Y. Pak, *Multiple stepwise pattern for potential of mean force in unfolding the thrombin binding aptamer in complex with Sr2+*. The Journal of chemical physics, 2011. **135**(22): p. 12B610.
324. Islam, B., et al., *Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale.* Nucleic acids research, 2013. **41**(4): p. 2723-2735.
325. Stadlbauer, P., et al., *Structural dynamics of possible late-stage intermediates in folding of quadruplex DNA studied by molecular simulations.* Nucleic acids research, 2013: p. gkt412.
326. Bergues-Pupo, A.E., et al., *Role of the central cations in the mechanical unfolding of DNA and RNA G-quadruplexes.* Nucleic acids research, 2015. **43**(15): p. 7638-7647.
327. Linak, M.C., R. Tourdot, and K.D. Dorfman, *Moving beyond Watson–Crick models of coarse grained DNA dynamics.* The Journal of chemical physics, 2011. **135**(20): p. 11B613.
328. Rebic, M., et al., *Multiscale simulations of human telomeric G-quadruplex DNA.* The Journal of Physical Chemistry B, 2014. **119**(1): p. 105-113.
329. Bergues-Pupo, A.E., et al., *Mesoscopic model for DNA G-quadruplex unfolding*. arXiv preprint arXiv:1705.10778, 2017.
330. Mergny, J.-L., A.-T. Phan, and L. Lacroix, *Following G-quartet formation by UV-spectroscopy.* FEBS letters, 1998. **435**(1): p. 74-78.

331. Long, X., et al., *Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy.* Nucleic acids research, 2013. **41**(4): p. 2746-2755.